الفصل الثالث

التحليل الإحصائي

برنامج SPSS

(Statistical Package for Social Sciences)

زيد ناصر عليمات
بكالوريوس فيزياء
دبلوم عالي وماجستير كمبيوتر
الأردن

قناة أكاديمية السندباد

➢ ٢,١٢,١ التعرف على النماذج
➢ ٢,١٢,٢ إنشاء نموذج باستخدام قالب صفحة النموذج
➢ ٢,١٢,٣ إنشاء نموذج باستخدام معالج صفحة النموذج
➢ ٢,١٢,٤ إنشاء نموذج من البداية

☑ ٢,١٣ **الأطر Frames** ١١٠
➢ ٢,١٣,١ التعرف على الأطر
➢ ٢,١٣,٢ إنشاء الأطر باستخدام قوالب FrontPage

☑ ٢,١٤ **نشر موقع الويب وإدارته** ١١٣
➢ ٢,١٤,١ نشر موقع Web
➢ ٢,١٤,٢ نشر الموقع محلياً على القرص الصلب
➢ ٢,١٤,٣ نشر الموقع على خادم Web (الإنترنت)
➢ ٢,١٤,٤ إدارة موقع Web
➢ ٢,١٤,٥ للحصول على تقارير عن موقع Web

الفصل الثالث: التحليل الإحصائي SPSS

☑ ٣,١ **مقدمة Introduction** ١١٨

☑ ٣,٢ **مقاييس النزعة المركزية Measure of Central Tendency** ١١٩
➢ ٣,٢,١ الوسط الحسابي (المعدل) The Arithmetic Mean
➢ ٣,٢,٢ الوسيط The Median
➢ ٣,٢,٣ المنوال أو القمة The Mode

☑ ٣,٣ **مقاييس التشتت Measures of Dispersion** ١٢٦
➢ ٣,٣,١ المدى Range
➢ ٣,٣,٢ الانحراف المتوسط The Mean Deviation
➢ ٣,٣,٣ التباين The Variance
➢ ٣,٣,٤ الانحراف المعياري The Standard Deviation
➢ ٣,٣,٥ معامل التغير Coefficient of Variation

☑ ٣,٤ **اختبار الفرضيات Hypothesis testing** ١٣١

☑ ٣,٥ **الارتبـــاط Correlation** ١٣٣
➢ ٣,٥,١ معامل ارتباط بيرسون الخطي (Pearson's Correlation Coefficient)
➢ ٣,٥,٢ معامل ارتباط سبيرمان للرتب (Spearman's Correlation Coefficient)

☑ ٣,٦ **الانحدار البسيط Linear Regression** ١٣٧

☑ ٣,٧ نوافذ برنامج SPSS...139
➢ ٣,٧,١ نافذة محرر البيانات Data Editor Window
➢ ٣,٧,٢ نافذة المخرجات Output Window
➢ ٣,٧,٣ نافذة التعليمات Syntax Windows
☑ ٣,٨ شريط الأدوات ١٤٠
☑ ٣,٩ القوائم الرئيسة ١٤١
➢ ٣,٩,١ فتح ملف بيانات مخزن Open
➢ ٣,٩,٢ حذف متغير (عمود Delete Variable (column
➢ ٣,٩,٣ حذف صف (حالة) Delete Row(case)
➢ ٣,٩,٤ نسخ و قص Copy & Cut
➢ ٣,٩,٥ البحث عن القيم Finding Values
☑ ٣,١٠ قائمة البيانات Data ١٤٤
➢ ٣,١٠,١ إدراج متغير (عمود)Insert Variable (column
➢ ٣,١٠,٢ إدراج صف (حالة) Insert cases (Row)
➢ ٣,١٠,٣ الذهاب إلى حالة ما Go To Case
➢ ٣,١٠,٤ ترتيب البيانات Sort Cases
➢ ٣,١٠,٥ تقسيم الملفات Split File
➢ ٣,١٠,٦ دمج الملفات Merge Files
➢ ٣,١٠,٧ اختيار الحالات Select Cases
☑ ٣,١١ قائمة التحويلات Transformation ١٥٠
➢ ٣,١١,١ الأمر Compute لإجراء العمليات الحسابية
➢ ٣,١١,٢ الأمر Count
➢ ٣,١١,٣ الأمر إعادة الترميز Recode
☑ ٣,١٢ قائمة Analysis ١٥٦
➢ Descriptive Statistics 3.12.1
▪ Frequencies 3.12.1.1
▪ Descriptive 3.12.1.2
➢ Explore 3.12.2
☑ ٣,١٣ اختبار الفرضيات ١٦١
➢ ٣,١٣,١ اختبار T
☑ ٣,١٤ الارتبـــاط Correlation ١٦٤
☑ ٣,١٥ الانحدار البسيط Linear Regression ١٦٥

الفصل الرابع: الاختراق والأمن Security

☑ ٤,١ المقدمة ١٦٩
➢ ٤,١,١ خصوصية المعلومات (Privacy)
➢ ٤,١,٢ سلامة المعلومات (Integrity)
➢ ٤,١,٣ التحقق من هوية الأطراف الأخرى (Authentication)
☑ ٤,٢ الاختراق ١٧٢
➢ ٤,٢,١ ما هو الاختراق.
➢ ٤,٢,٢ أسباب الاختراق ودوافعه
➢ ٤,٢,٣ أنواع الاختراق
➢ ٤,٢,٤ أثار الاختراق
➢ ٤,٢,٥ اختراق المواقع
➢ ٤,٢,٦ كيف تواجه الاختراق؟
☑ ٤,٣ مخترقي الأجهزة والشبكات (الهاكرز) ١٧٦
➢ ٤,٣,١ تاريخ الهاكرز
➢ ٤,٣,٢ برامج الهاكرز
➢ ٤,٣,٣ ميكانيكية الاختراق
☑ ٤,٤ الفيروسات Viruses ١٧٩
➢ ٤,٤,١ ما هو الفيروس؟.
➢ ٤,٤,٢ نشأة وتاريخ الفيروسات.
➢ ٤,٤,٣ انتشار الفيروسات و طرق العدوى.
➢ ٤,٤,٤ كيف تعمل الفيروسات.
➢ ٤,٤,٥ أنواع الفيروسات.
➢ ٤,٤,٦ علاج الفيروسات و التخلص منها.
➢ ٤,٤,٧ البرامج الخبيثة
➢ ٤,٤,٨ الإنترنت و طرق الوقاية
☑ ٤,٥ التشفير Encryption ١٨٥
➢ ٤,٥,١ مقدمة.
➢ ٤,٥,٢ ما هو التشفير Encryption.
➢ ٤,٥,٣ التشفير المتماثِل (Symmetric Cryptography
➢ ٤,٥,٤ التشفير غير المتماثل (Asymmetric Cryptography
☑ ٤,٦ البصمة الإلكترونية للرسالة والتوقيع الرقمي.. ١٨٧
➢ ٤,٦,١ البصمة الإلكترونية (Message Digest)
➢ ٤,٦,٢ التوقيع الرقمي (Digital Signature)

☑ ٣,١ مقدمة Introduction :

يلعب علم الإحصاء دوراً هاماً في مختلف المجالات الاقتصادية والتربوية والاجتماعية، وقد بذل الباحثون جهداً عظيما في عملية التحليل الإحصائي وذلك لصعوبة العمليات الحسابية آنذاك، أدى ذلك إلى التقليل من حجم العينة مما أدى إلى ظهور مشكلة وهي أن هذه العينة لا تفي بالغرض وكانت النتائج غير دقيقة ومن الخطاء أن تعمم على مجتمع الدراسة.

تمكن علم الإحصاء من الاستفادة من الثورة التقنية وتقدم علم الحاسب، مما مكن الباحثون من زيادة حجم العينة بدون أيـــة صعوبات، وقربنا ذلك إلى النتائج الدقيقة، وظهرت برامج كثيرة في هذا المجال منها برنامج التحليل الإحصائي الشهير SPSS، فكان أداة مفيدة لطلبة الجامعات والعاملين في مجال الأبحاث على حد سواء.

في هذا الفصل سنستعرض مبادئ الإحصاء مثل:

- مقاييس النزعة المركزية Measure of Central Tendency
- مقاييس التشتت Measures of Dispersion
- اختبار الفرضيات Hypothesis testing
- الارتبــــاط Correlation
- الانحدار البسيط Linear Regression

ثم ننتقل إلى التحليل الإحصائي بمختلف أساليبه ومراحله، منتقلين من مرحلة المعالجة إلى مرحلة تحليل النتائج.

☑ ٣,٢ مقاييس النزعة المركزية Measure of Central Tendency

انه من الطبيعي أن تتركز اكثر القيم في المنتصف كالأطوال والأوزان والأعمار وغيرها من ظواهر. ومقاييس النزعة المركزية تبحث في تقدير قيمة تتمركز حولها أغلبية البيانات وأهم هذه المقاييس:

1. *الوسط الحسابي The Arithmetic Mean*
2. *الوسيط The Median*
3. *المنوال The Mode*

➢ ٣,٢,١ الوسط الحسابي (المعدل): The Arithmetic Mean

الوسط الحسابي هي القيمة الناتجة من قسمة مجموع قيم المتغير على عددها ويرمز له بالرمز $\bar{y}$.

طرق حسابه:

أ. حالة البيانات غير المبوبة:

إن كان لديك n من القيم أو المشاهدات $y_1, y_2, \ldots\ldots\ldots\ldots y_n$ فإن الوسط الحسابي لها هو:

$$\bar{y} = \frac{\sum y_i}{n}$$

مثال ١: أحسب الوسط الحسابي لأطوال مجموعة من الأطفال مبينة في جدول التالي. *جدول ٣,١*

٨١	٧١	٦٧	٤٠	٩٥	٩٠	٦٣	٦٢
٧٨	٨٤	٩٣	٩١	٧٠	٤٨	٧٥	٨٢
٧٣	٧١	٨١	٥٦	٦٥	٩٢	٧٠	٧٤
٨٦	٧٢	٩٣	٦٥	٥١	٨٥	٦٨	٨٣
٣٥	٨٦	٤٣	٧٤	٧٣	٨٣	٩٠	٦٨
٧٥	٦٧	٧٢	٩٠	٧١	٧٦	٩٢	٩٣
٩٠	٨٨	٩١	٩٧	٧٢	٦١	٨٠	٨٠
٦٦	٧٨	٥٩	٨٠	٧٤	٩٩	٧٠	٧٤
٦٣	٧٩	٣٠	٧٦	٨٢	٨٣	٦٣	٦٠
٨٠	٨٩	٨٨	٧٧	٩٨	٨٨	٨١	٧٩

الحل :

$$\bar{y} = \frac{\sum y_i}{n} = \frac{80 + 84 + \ldots\ldots\ldots + 75}{80}$$

$$= \frac{6046}{80} = 75.58\ cm$$

أي أن معدل طول الأطفال هو ٧٥,٥٨سم

ب- البيانات المبوبة (فئات):

إذا كانت البيانات مرتبة في فئات فالوسط الحسابي هو =

$$\bar{y} = \frac{\sum_{i=1}^{k} f_i y_i}{\sum_{i=1}^{k} fLi}$$

حيث $y_1, y_2 \ldots\ldots\ldots y_k$. تمثل مراكز الفئات في جدول التوزيع التكراري و $f_1, f_2 \ldots\ldots\ldots f_k$. تكراراتها .

خطوات إيجاد الوسط الحسابي في بيانات مبوبة هي كالآتي:

(١) نشكل الفئات : نجد المدى = اكبر قيمة – اصغر قيمة

(٢) نقسم المدى على طول الفئة وليكن طولها ٩ وحدات.

(٣) ثم عيّن مراكز الفئات حيث مركز الفئة = (الحد الأدنى للفئة + الحد الأعلى) /٢.

(٤) ضرب مركز كل فئة بمقدار تكرارها.($f_i\ y_i$).

(٥) قسمة مجموع [حاصل ضرب مركز الفئة × تكرارها] على مجموع التكرارات.

مثال ٢: استخرج الوسط الحسابي لأطول الأطفال من البيانات في المثال السابق

الحل:

عيّن مركز الفئات ثم نضرب مركز كل فئة في تكرارها كما في الجدول التالي:

جدول ٣,٢

الفئات	التكرار	مركز الفئات	*التكرار × مركز الفئات*
٣١-٤٠	٣	٣٥,٥	١٠٦,٥
٤١-٥٠	٢	٤٥,٥	٩١
٥١-٦٠	٥	٥٥,٥	٢٧٧,٥
٦١-٧٠	١٤	٦٥,٥	٩١٧
٧١-٨٠	٢٤	٧٥,٥	١٨١٢
٨١-٩٠	٢٠	٨٥,٥	١٧١٠
٩١-١٠٠	١٢	٩٥,٥	١١٤٦
	$\sum f_{i=80}$		$\sum f_i x_i = 6060$

أي أن معدل طول الأطفال هو ٧٥,٧٥سم

لاحظ بأن هذا الرقم يختلف قليلاً عن الوسط الحسابي (٧٥,٥٨سم) لنفس البيانات قبل تبويبها ووضعها في جدول توزيع تكراري (٧٥,٧٥سم).، وهذا الفرق بين الرقمين يعود إلى فقدان المعلومات بسبب وضع المفردات في فئات فنحن نفرض بأن طول كل الأطفال في فئة معينة مساوياً لمركز تلك الفئة.

خواص الوسط الحسابي:

- عند إضافة عدد ثابت (k) إلى كل قيمة من قيم المشاهدات فإن الوسط الحساب للقيم الجديدة = الوسط الحسابي للقيم الأصلية + العدد الثابت (k) أي أن $\bar{x}_i = \bar{y}_i + k$

مثال ٣: لنفرض أن لديك القيم التالية: $y_i = 10, 4, 3, 13, 10$ فالوسط الحسابي لها هو: $\bar{y} = \frac{\sum y_i}{n} = \frac{40}{5} = 8$

فإذا أضفنا لكل من هذه القيم قيمة ثابتة ولتكن ٤

فالقيم الجديدة ستصبح: $x_i = 14, 8, 7, 17, 14$

والوسط الحسابي للقيم الجديدة هو: $\bar{x} = \frac{\sum x_i}{n} = \frac{60}{5} = 12$

الذي هو في الحقيقة: $\bar{x} = \bar{y} + 3$
$= 8 + 4 = 12$

- *مجموع انحرافات القيم عن وسطها الحسابي تساوي صفراً*

أي (للبيانات غير المبوبة) $\sum(y_i - \bar{y}) = 0$ أو (للبيانات المبوبة) $\sum f_i(y_i - \bar{y}) = 0$

والجدول التالي يوضح ذلك:

جدول ٣,٣

y_i	$(y_i - \bar{y})$
١٠	2
٤	-4
٣	-5
١٣	5
10	2
$\sum y_i = 40$	$\sum(y_i - \bar{y}) = 0$
$\bar{y} = 8$	

- إذا ضربت كل قيمة من قيم المشاهدات في قيمة ثابتة (k) فإن: الوسط الحسابي للقيم الجديدة = الوسط الحسابي للقيم الأصلية × العدد الثابت k $\bar{z} = k\bar{y}$

مثال ٤: لديك القيم التالية: $y_i = 10, 4, 3, 13, 10$ نجد أن $\bar{y} = 8$ فإذا كان $z_i = 5y_i$ أوجد قيمة $\bar{z}$

الحل:

$$\therefore z_i = 50, 20, 15, 65, 50$$

$$\therefore \bar{z} = \frac{\sum z_i}{n} = \frac{200}{5} = 40$$

أو

$$\therefore \bar{z} = (5)(\bar{y}) = 5(8) = 40$$

- الوسط الحسابي لمجموع قيم متغيرين = مجموع الوسطين الحسابيين للمتغيرين، أي:

مثال ٥: ليكن لديك العينتين التاليتين أوجد الوسط الحسابي لـــ z.

جدول ٣,٤

y_i	x_i	$zi = xi + yi$
٨	٤	١٢
٥	١٠	١٥
٤	٣	٧
٢	١٣	١٥
٤	١٠	١٤
$\bar{x}$ =٤,٦	$\bar{y}$ =٨	$\bar{z}$ = ١٢,٦

من هنا يتضح بأن:

$$\bar{z} = \bar{x} + \bar{y}$$
$$= 4.6 + 8 = 12.6$$

٣,٢,٢ الوسيط The Median

(أ) البيانات غير المبوبة:

إذا كان لديك n من القيم أو المشاهدات $y_1, y_2 ...., y_n$ ورتبت ترتيباً تصاعدياً (أو تنازلياً)

(١) فإذا كانت n عدد فردي فإن الوسيط هو القيمة التي ترتيبها $\frac{n+1}{2}$ (ويرمز له بـــ $\overline{Me}$) فالوسيط هو $\overline{Me} = y_{(n+1)/2}$

(٢) أما إذا كانت n عدد زوجي فإن الوسيط هو الوسط الحسابي للقيمتين اللتين ترتيبهما $\frac{n}{2}+1, \frac{n}{2}$

أي: $\overline{Me} = \frac{y_{n/2} + y_{(n/2)+1}}{2}$

مثال ٦ : أوجد الوسيط لدرجات طلاب في امتحان مادة الفيزياء إذا كانت الدرجات هي: ٨٨، ٩٠، ٧٦، ٧٩، ٧٥

الحل: نرتب الدرجات تصاعدياً ٧٥، ٧٦، ٧٩، ٨٨، ٩٠ وبما أن عدد الأرقام فردي (n=5)

إذن فالوسيط هو القيم التي ترتيبها: $3 = \frac{5+1}{2} = \frac{n+1}{2}$

$\overline{Me} = y_3 = 79$ أي الوسيط = ٧٩

مثال ٧ : أوجد الوسيط للقيم التالية: $y_i = 2,7,8,5,1,9,10,6$ نرتب القيم تصاعدياً: $y_i = 1, 2, 5, 6, 7, 8, 9, 10$ وبما أن عدد القيم هو زوجي (n=8) إذن فالوسيط الحسابي للقيمتين اللتين ترتيبهما: (n/2)+1, (n/2)

الموقع رقم $4 = \frac{8}{2} = (n/2)$ و (n/2)+1=5

$$\therefore \overline{Me} = \frac{y_4 + y_5}{2} = \frac{6+7}{2} = 6.5$$

(ب) البيانات المبوبة:

إذا كانت $y1, y2 ...., yk$ تمثل مراكز الفئات في جدول التوزيع التكراري مع تكراراتها: $f1, f2 ...., fk$ على التوالي. فقيمة الوسيط لهذه البيانات (بالاستعانة بجدول التوزيع التكراري المتجمع الصاعد) هو

$$\overline{Me} = L_1 + \left[\frac{(\sum f_{i/2}) - F_i}{f_i}\right] w$$

حيث أن:- L_1 : الحد الأدنى الحقيقي لفئة الوسيط.

$\sum f_i$: مجموع التكرارات.

Fi : التكرار المتجمع عند بداية فئة الوسيط.

f : تكرار فئة الوسيط.

W: طول فئة الوسيط = التكرار المتجمع عند نهاية فئة الوسيط–التكرار المتجمع عند بداية فئة الوسيط.

خطوات إيجاد الوسيط هي:

١) عمل جدول توزيع تكراري تجميعي تصاعدي (أو تنازلي)

٢) إيجاد ترتيب الوسيط وهو $\frac{\sum f_i}{2}$

٣) تحدد فئة الوسيط وهي الفئة التي تقع قيمة الوسيط بين حديها وذلك عن طريق إيجاد قيمتين متتاليتين في التكرار التجميعي التصاعدي يقع بينهما ترتيب الوسيط. يقابل هاتين القيمتين حدا فئة الوسيط الأدنى والأعلى ويستحسن أخذ الحدود الحقيقية لهذه الفئة. ثم نطبق القانون.

ملاحظة: من الممكن إيجاد ترتيب الوسيط بـ $\left(\frac{\sum f_i}{2}\right)$ إذا كان عدد المفردات فردياً أو بـ $\left(\frac{\sum f_i}{2}+1\right)$ إذا كان عدد المفردات زوجياً. ولكن نظراً لكون المفردات كبيراً في التوزيعات التكرارية فتستخدم $\left(\frac{\sum f_i}{2}\right)$ لإيجاد ترتيب الوسيط.

مثال ٨: أوجد الوسيط للتوزيع التكراري في جدول

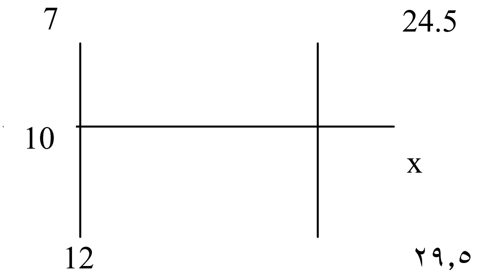
الحل: سوف نستخدم أولا طريقة النسبة والتناسب *جدول ٣,٥*

الفئات	f_i	أقل من	التكرار التراكمي
١٠-١٤	١	أقل من ١٤,٥	١
١٥-١٩	٢	أقل من ١٩,٥	٣
٢٠-٢٤	٤	أقل من ٢٤,٥	٧
٢٥-٢٩	٥	أقل من ٢٩,٥	١٢
٣٠-٣٤	٥	أقل من ٣٤,٥	١٧
٣٥-٣٩	٣	أقل من ٣٩,٥	٢٠
	٢٠		

١٠ فئة الوسيط

ترتيب الوسيط $\frac{\sum f_i}{2}=\frac{20}{2}=10$

أي أن قيمة الوسيط هو طول العينة التي ترتيبها ١٠ (بعد ترتيب القيم تصاعدياً أو تنازليا). وفي جدول التوزيع التكراري التجميعي التصاعدي نرى بأن ١٠ هي واقعة بين الرقمين ٧، ١٢. إذن فئة الوسيط هي:



المدى	التكرار
(٢٩,٥ -٢٤,٥)	(١٢-٧)
X	(١٠-٧)

ثم نقوم بالضرب التبادلي

$$x=\frac{5\times 3}{5}=3$$

$$\therefore Me = 24.5+3=27.5$$

والآن سنستخدم العلاقة السابقة لإيجاد الوسيط حيث قيم العلاقة كما يلي:

$L_1 = 24.5$: الحد الأدنى الحقيقي لفئة الوسيط.

$F_i = 7$: التكرار المتجمع عند بداية فئة الوسيط.

$f_i = 4$: تكرار فئة الوسيط.

W = 10 – 14 = 4 : طول فئة الوسيط

$$= 24.5 + \left[\frac{10 - 7}{4}\right](4)$$

$= 27.5$ وهي نفس النتيجة السابقة

➢ ٣,٢,٣ المنوال أو القمة The Mode

(أ)البيانات غير المبوبة:

تعريف:إذا كان لديك n من المشاهدات $y1, \ y2 \ldots\ldots\ldots, yn$.. فإن المنوال لهذه المشاهدات هي المشاهدة الأكثر تكراراً بين هذه المشاهدات ويرمز له بـ $\overline{Mo}$ ومن هذا يتضح بأنه قد يكون هناك منوالاً واحداً لهذه المشاهدات وعندها يسمى التوزيع وحيد القيمة ، أو يكون لها منوالان وعندها يسمى التوزيع ذو قيمتين وقد يكون لها أكثر من منوالين أو قد لا يوجد لهذه المشاهدات أي منوال.

مثال ٩ : أوجد المنوال لكل من البيانات التالية:

(أ) 2,6,3,5,5,9,5,8,6,2

(ب) 50, 44, 51, 48, 49

الحل:

(أ) المفردة ٥ هي أكثر المفردات تكراراً فهي المنوال $\therefore \overline{Mo} = 5$

(ب) لا يوجد منوال لهذه المفردات.

(ب)البيانات المبوبة :

إذا كانت القيم $y_1, y_2 \ldots\ldots\ldots\ldots\ldots, y_k$ تمثل مراكز الفئات في جدول التوزيع التكراري مع تكراراتها $f_1, f_2 \ldots\ldots\ldots\ldots\ldots f_k$ على التوالي. فإن المنوال:

$$\overline{Mo} = L_1 + \left(\frac{d_1}{d_1 + d_2}\right) w$$

حيث أن:

فئة المنوال: تلك الفئة التي تملك أكبر التكرارات وأن:

L_1: الحد الأعلى الحقيقي لفئة المنوال.

d_1: الفرق بين تكرار فئة المنوال والفئة السابقة لها.

d_2: الفرق بين فئة المنوال والفئة اللاحقة لها.

w: طول الفترة

الفئات	التكرار
١٠ – ١٤	١١
١٥ – ١٩	٢١
٢٠ – ٢٤	٤٠
٢٥ – ٢٩	٥١
٣٠ – ٣٤	٤٠

	١٦٣

مثال ١٠: أوجد المنوال للجدول التوزيع التكراري التالي.

الحل: إن الفئة (٢٥ - ٢٩) لها أكبر التكرارات (٥١) فهي إذن فئة المنوال

جدول ٣,٦

الحد الأدنى الحقيقي لفئة المنوال

$L_1=24.5$

$d_1 = 51 - 40 = 11$

$d_2 = 51 - 40 = 11$

$w = 4$

$$\overline{Mo} = 24.5 + \left(\frac{11}{11+11}\right)(4) = 26.5$$

مقاييس أخرى للتوسط أو التمركز:

(١) الربيع الأدنى والربيع الأعلى: Upper and Lower Quartiles

الربيع الأدنى (أو الربيع الأول): هو قيمة المفردة التي تقسم المفردات (بعد ترتيبها تصاعدياً أو تنازلياً) إلى قسمين بحيث يسبقها ٢٥% من المفردات ويليها ٧٥% من المفردات.

والربيع الأعلى (أو الربيع الثالث): هو قيمة المفردة التي تقسم المفردات (بعد ترتيبها تصاعدياً أو تنازلياً) إلى قسمين بحيث يسبقها ٧٥% من المفردات ويليها ٢٥% من المفردات.

ملاحظة: يمكن اعتبار الوسيط بالربيع الأوسط (أو الربيع الثاني) لأنه يتوسط المفردات أي يقسم المفردات إلى قسمين (بعد ترتيبها تصاعدياً أو تنازلياً) بحيث يسبقه ٥٠% ويليه ٥٠% من المفردات.

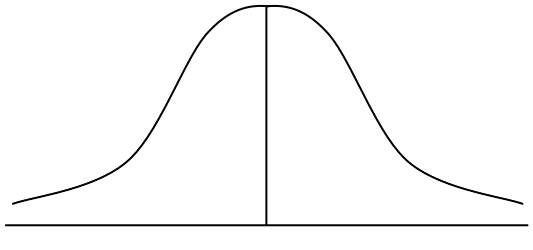
(٢) العشريات Deciles والمئينيات Percentiles

العشريات والمئينيات هي مواقع أيضاً على التوزيع تقسم المفردات (بعد ترتيبها تصاعدياً أو تنازلياً) فمثلاً العشير الثالث: هو قيمة المفردة أو المشاهدة التي يسبقها $\frac{3}{10}$ ويليها $\frac{7}{10}$ المفردات..

أما المئين السبعون فهو قيمة المفردة التي يسبقها ٧٠% من المفردات ويليها ٣٠% من المفردات.

- العلاقة بين بعض المتوسطات للتوزيعات ذات منوال واحد

(١) إذا كان التوزيع متماثلاً Symmetrical فإن قيمة الوسط الحسابي والوسيط والمنوال تتساوى تماماً

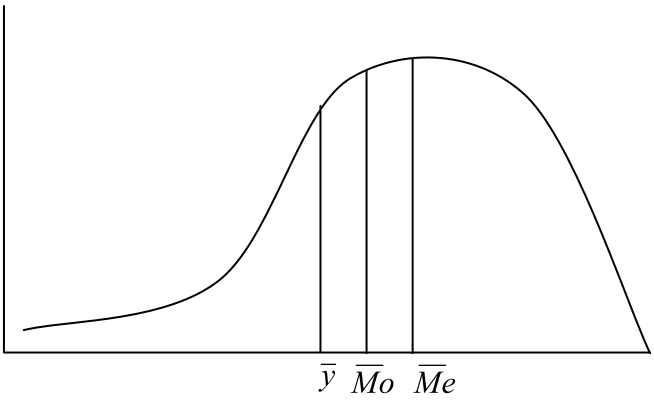


منحني متماثل

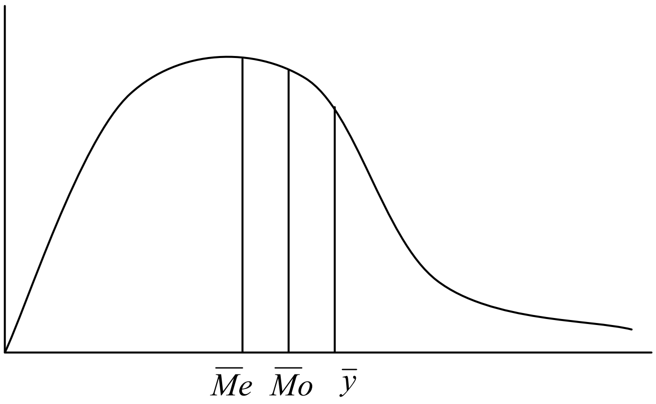
(٢) أما إذا كان التوزيع غير متماثل Asymmetrical (ملتوى التواء معتدلاً) فإن:

الوسط الحسابي – المنوال = ٣ (الوسط الحسابي – الوسيط) $\bar{y} - \bar{M}o = 3(3\,\bar{y} - \bar{M}e)$

التواء سالب



التواء موجب



☑ ٣,٣ مقاييس التشتت Measures of Dispersion

تعطي مقاييس التشتت مدى الاختلاف بين عناصر أفراد الدراسة و صورة عن مدى الانسجام بين مشاهدات الدراسة ، وقد تعطي مقاييس التشتت في بعض الحالات صورة افضل عن نتائج الدراسة مقارنة مع ما تعطيه مقاييس النزعة المركزية .ولتوضيح ذلك نورد المثال التالي:

مثال ١١: إذا كان لديك دولتين وكان الدخل الشهري لثلاثة عمال من كل دولة كما يلي:

الدولة الأولى :	١٥٠٠	٣٥٠	٢٥٠
الدولة الثانية :	٧٠٠	٦٥٠	٧٥٠

نلاحظ إن متوسط الدخل الشهري في الدولتين متساوي حيث يبلغ ٧٠٠ دينار ، الاّ أننا نلاحظ أن التباين في دخل العاملين في الدولة الأولى يبلغ ١٢٥٠ (١٥٠٠ – ٢٥٠) ، في حين يبلغ ١٠٠ فقط في الدولة الثانية (٧٥٠ – ٦٥٠) . نلاحظ أن الرواتب منخفضة في الدولة الأولى باستثناء الراتب الأولى الذي يعتبر عالي مقارنة ببقية الرواتب أما الدولة الثانية فالرواتب افضل. وعلى الرغم من هذه الحقيقة قد تلجأ الدولة الأولى إلى استخدام المتوسط الحسابي للتظاهر بارتفاع الدخل الشهري، ويكون في ذلك تضليل لمن لا يعرف هذه الحقيقة . وبالتالي فان استخدام مقاييس التشتت قد تعطي صورة أوضح عن مثل هذه الحالة وغيرها من الحالات المشابهة.

من أهم مقاييس التشتت ما يلي :

- *المدى* *Range*
- *الانحراف المتوسط* *The Mean Deviation*
- *التباين* *The Variance*

- الانحراف المعياري   *The Standard Deviation*
- معامل التغير   *Coefficient of Variation*

٣,٣,١ المدى **Range**

تعريف : هو الفرق بين أعلى قيمة و أقل قيمة من بيانات الدراسة. ويعّرف المدى في البيانات المجمعة (التوزيع التكراري) بأنه الفرق بين الحد الأعلى للفئة العليا والحد الأدنى للفئة الدنيا.

$$Range = Max - Min$$

مثال ١٢: إذا كان أطوال طلاب كما يلي :   ٩٠ ، ١١٠ ، ١١٥ ، ١٢٠ ، ١٢٥ ، ١٤٠

فان المدى = ١٤٠ − ٩٠ = ٥٠

أي أن القيم متباعدة عن بعضها البعض بالرغم من أن القيم واقعة بين ٩٠ وبين ١٤٠.

ملاحظة: المدى لا يعتمد على جميع البيانات و إنما يعتمد على اكبر قيمة واقل قيمة ، لهذا فهو يتأثر بالقيم الشاذة ويعتبر مقياس ضعيف و لحل هذه المشكلة نحذف أعلى قيمة وأصغر قيمة من البيانات وبالتالي نحسّن المدى.

في المثال السابق نحذف القيمتين ١٤٠ و ٩٠ فيكون

المدى = ١٢٥ − ١١٠ = ١٥   أي أن البيانات متقاربة ومتجانسة

3.3.2 الانحراف المتوسط **Mean Deviation**

تعريف: هو مجموع انحرافات البيانات عن وسطها الحسابي ولكن هذا المجموع يساوي صفراً لان مجموع الانحرافات الموجبة عن الوسط الحسابي يساوي مجموع الانحرافات السالبة ، لذلك يجب حذف الإشارة السالبة لنحصل على مقياس ذي معنى. أن إحدى الطرق لتخلص من الإشارة السالبة هي بأخذ القيمة المطلقة *(Absolute Value)* التي تعرف بالمعادلة :

$$|X| = \begin{cases} X , & X \geq 0 \\ -X , & X < 0 \end{cases}$$

وباستعمال القيمة المطلقة نحصل على تعريف الانحراف المتوسط

إذا كان لديك البيانات $X_1, X_2, \ldots\ldots, X_n$

فإن انحرافها المتوسط يكون

$$M.D = \frac{\sum_{i=1}^{n} |X_i - \bar{X}|}{n}$$

حيث $\bar{X}$ الوسط الحسابي للبيانات

مثال ١٣: أوجد الانحراف المتوسط للبيانات التالية: ٥، ٦، ٩، ١٠، ٤، ٢

$$\bar{X} = \frac{\sum X_i}{n} = \frac{5 + 6 + 9}{}$$

$$M.D = \frac{\sum_{i=1}^{n} |X_i - \bar{X}|}{n}$$

$$= \frac{|5-6| + |6-6| + |9-6| + |10-6| + |4-6| + |2-6|}{6}$$

$$= \frac{1 + 0 + 3 + 4 + 2 + 4}{6} = 2\tfrac{1}{3}$$

أما في حالة البيانات المجمعة (التوزيعات التكرارية) فنستعمل التعريف التالي :

تعريف : إذا كانت مراكز التوزيع التكراري $X_1, X_2, ...,X_n$ *وكانت التكرارات المقابلة لها* $f_1, f_2, ..., f_h$

فإن الانحراف المتوسط يكون على الشكل

$$M.D = \frac{\sum_{i=1}^{h} \left| X_i - \overline{X} \right| . f_i}{n}$$

حيث $\overline{X}$ *= الوسط الحسابي ، و* n *= مجموع التكرارات.*

مثال ١٤ : أحسب الانحراف المتوسط للتوزيع التكراري في الجدول 3.7

جدول ٣,٧

X_i	f_i
4	3
6	5
8	7
10	9
12	2

الحل : نرتب الحسابات في الجدول 1.2

جدول ٣,٨

X_i	f_i	Xf_i	$X_i - \overline{X}$	$\left\|X_i - \overline{X}\right\|$	$\left\|X_i - \overline{X}\right\| f_i$
4	3	12	-4.15	4.15	12.46
6	5	30	-2.15	2.15	10.77
8	7	56	-0.15	0.15	1.08
10	9	90	1.85	1.85	16.62
12	2	24	3.85	3.85	7.69
المجموع	26	212			48.26

الوسط الحسابي $\overline{X} = 212 / 26 = 8.15$

$$M.D = \frac{\sum_{i=1}^{h} \left| X_i - \overline{X} \right| . f_i}{n}$$

$$M.D = 48.26 / 26 = 1.87$$

الانحراف المتوسط يعتمد على جميع مفردات البيانات وهو سهل الحساب الاّ انه لا يخضع للعمليات الجبرية بسهولة ، ويجب معرفة جميع المفردات إذا أردت معرفة الانحراف المتوسط للمجموعة الناتجة من دمج مجموعتين من البيانات.

➢ ٣,٣,٣ التباين The Variance

تعريف : هو مجموع مربع انحرافات القيم عن الوسط الحسابي مقسوما على عدد القيم ، وهي الطريقة الثانية للتخلص من الإشارة السالبة لانحرافات القيم عن الوسط الحسابي ، بمقارنة التباين والمدى نلاحظ أن التباين يمتاز على المدى في انه اقل تأثرا بالقيم الشاذة ويأخذ بالاعتبار عند احتسابه جميع قيم المشاهدات.

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{(n-1)}$$

مثال ١٥: احسب التباين للبيانات 3, 6, 8, 2, 1, 4

الحل : الوسط الحسابي $\overline{X} = (3 + 6 + 8 + 2 + 1 + 4) / 6 = 4$

انحرافات البيانات عن الوسط الحسابي هي (3-4), (6-4), (8-4), (2-4), (1-4), (4-4)

ومربع هذه الانحرافات 1, 4, 16, 4, 9, 0

$$S^2 = \frac{1+4+16+4+9+0}{5} = \frac{34}{5} = 6.8$$

تعريف : إذا كانت مراكز التوزيع التكراري هي $X_1, X_2, \ldots, X_n$

وكانت التكرارات المقابلة لها $f_1, f_2, \ldots, f_h$ *فان*

$$S^2 = \frac{\sum_{i=1}^{h} (X_i - \overline{X})^2 . f_i}{n-1}$$

مثال ١٦: *احسب التباين للتوزيع التكراري في المثال ١٤.*

الحل : نرتب خطوات الحل في الجدول ٣,٩

جدول ٣,٩

X_i	f_i	Xf_i	$(X_i - \overline{X})$	$(X_i - \overline{X})^2$	$(X_i - \overline{X})^2 . f_i$

4	3	12	-4.15	17.25	51.76
6	5	30	-2.15	4.64	23.20
8	7	56	-0.15	0.02	0.17
10	9	90	1.85	3.41	30.67
12	2	24	3.85	14.79	29.59
المجموع	26	212			135.38

الوسط الحسابي $\overline{X} = 212 / 26 = 8.15$

والتباين يساوي $S^2 = 135.38 / 25 = 5.42$

➢ 3.3.4 الانحراف المعياري Standard Deviation

تعريف : هو الجذر التربيعي للتباين أي أن

للبيانات الأولية:

$$S = \sqrt{S^2} = \sqrt{\frac{\sum (X_i - \overline{X})^2}{n-1}}$$

للبيانات المجمعة (التوزيع التكراري ذي الفئات):

$$S = \sqrt{S^2} = \sqrt{\frac{\sum (X_i - \overline{X})^2 . f_i}{n-1}}$$

أن طريقة حساب التباين طويلة وتحتاج إلى حساب أرقام كثيرة، وخاصة إذا احتوى الوسط الحسابي على كسور ، لهذا تم إيجاد طريقة مختصرة تجرى بالآلة الحاسبة وهي :

للبيانات الأولية

$$S^2 = \frac{1}{n-1}[\sum X_i^2 - n\overline{X}^2]$$

للبيانات المجمعة في توزيع تكراري ذي فئات

$$S^2 = \frac{1}{n-1}[\sum X_i^2 . f_i - n\overline{X}^2]$$

حيث n مجموع التكرارات

مثال 17: احسب التباين والانحراف المعياري للتوزيع التكراري في المثال ١٦ .

الحل : نرتب خطوات الحل في الجدول 1.4 جدول ٣,١٠

X_i	f_i	Xf_i	X_i^2	$X_i^2 f_i$
4	3	12	16	48
6	5	30	36	180
8	7	56	64	448
10	9	90	100	900
12	2	24	144	288
المجموع	26	212		1864

وبالتعويض في قانون التباين نجد

$$S^2 = \frac{1}{25}[1864 - 26 \times \left(\frac{212}{26}\right)^2]$$

$$S^2 = \frac{135.4}{25} = 5.42$$

والانحراف المعياري $S = \sqrt{5.42} = 2.33$

إذا كان تباين مجموعة من البيانات S_1^2 وعدد مفرداتها N1 وتباين مجموعة أخرى من البيانات S_2^2 وعدد مفرداتها N_2 فـان تباين المجموعة الكلية والناتج من دمجهما هو

$$S^2 = \frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}$$

مثال 18: أجرى طالب 8 تجارب لتعيين معامل المرونة فكان التباين في قياساته 0.٣ ثم أجرى طالب آخر ٩ تجارب فكان التباين في قياساته 0.2 ما هو تباين القياسات لجميع التجارب ؟

الحل:

$N_1 = 8$ $S_1^2 = 0.3$

$N_2 = 9$ $S_2^2 = 0.2$

$$S^2 = \frac{7 \times 0.3 + 8 \times 0.2}{15}$$

$$S^2 = \frac{3.7}{15} = 0.25$$

➢ ٣,٣,٥ معامل التغير Coefficient of Variation

$$C.V = \frac{S}{\bar{X}} \times 100\%$$

مثال 19: حدد أيهما اكثر تغيراً في البيانات في المثالين (15) و (16)

الحل:

المجموعة الأولى مثال 15

$$C.V = \frac{\sqrt{6.8}}{4} \times 100\% = 0.56 \times 100\%$$

$$C.V = 38.07\%$$

المجموعة الثانية مثال 16

$$C.V = \frac{\sqrt{5.42}}{8.15} \times 100\% = 0.29 \times 100\%$$

$$C.V = 29\%$$

ومنه نستنتج أن التغير في المثال ١٥ اكبر منه في المثال ١٦.

☑ ٣,٤ اختبار الفرضيات Hypothesis testing

<u>فرضية البحث Hypothesis</u>: عبارة عن تخمين مؤقت واستنتاج يتوصل إليه الباحث لحل المشكلة من خلال الإجابة عن السؤال الذي تمثله مشكلة البحث ويعبر عن رأي الباحث المبدئي ، فليست تفسير عشوائي، وإنما مستند إلى بعض المعلومات والخبـرة يتمسك بها الباحث حتى نهاية البحث، وعندها يتأكد من صحة الفرضيات أو بطلانها.

ملاحظات يجب على الباحث الانتباه إليها عند صياغته للفرضيات:

١. من الممكن أن يوجد في البحث فرضية واحدة رئيسية ، أو اكثر موزعة على جوانب البحث المختلفة. ولكن يجب أن تغطى كل جوانب البحث وتعطى التفسيرات الكافية لمشكلة البحث.
٢. من الممكن أن تصاغ الفرضية بالإثبات مثال "توجد علاقة قوية بين المستوى التعليمي للوالدين وعدد أفـراد العائلـة" أو أن تصاغ بالنفي، مثل "لا توجد علاقة قوية بين المستوى التعليمي... ".
٣. يجب أن تكون الفرضية قصيرة وبسيطة بحيث يسهل فهمها والتعرف على المتغير المستقل والمتغير التابع فيها.
٤. تشتمل الفرضية الواحدة عادة على متغير مستقل وآخر تابع، يؤثر الأول (المستقل) بالثاني (التابع).
٥. يمكن إثبات صحة الفرضية في نهاية البحث أو نفيها وذلك بالتحليل الإحصائي للمعلومات المجمعة من الاستبيان ، المقابلة ، الملاحظة أو أي طريقة أخرى.
٦. الفرضية ركن أساسي في كل أنواع البحوث وحتى البحوث ذات المنهج التاريخي.
٧. بعد إثبات الفرضية تصبح حقيقة، لأنها اختبرت وتم العثور على الدلائل التي تثبت صحتها. والحقيقة بعد تأكيدها وبلورتها بشكل اكثر استقراراً قد تتحول إلى نظرية. والنظرية قد تصبح قانون.

خصائص الفرضيات الجيدة

١. معقولية الفرضيات ، أي أن تكون منسجمة مع الحقائق العلمية المعروفة وان لا تكون خيالية أو مستحيلة أو متناقضة.
٢. قابلة للاختبار والقياس لهذا يجب على الباحث اتخاذ خطوات وإجراءات للتحقق من صحة الفروض.
٣. تفسر الظواهر المدروسة أي أن لديها القدرة على تقديم تفسير شامل لمشكلة البحث.
٤. بساطة ووضوح الفرضيات ، والابتعاد عن التعقيدات في صياغة الفروض باستخدام ألفاظ سهلة غير غامضة.
٥. يجب أن تكون محددة بدقة بعيدة عن العموميات، والعلاقة بين المتغيرات واضحة.
٦. أن تكون بعيدة عن التحيز الشخصي للباحث

أنواع الفرضيات: هناك نوعان من الفرضيات هما:
١. الفرضية العدمية (الصفرية) Null Hypothesis : H_O : وهي على شكل علاقة إيجابية بين المتغيرين المستقل والتابع مثل: "هناك علاقة إيجابية بين تزويد المدارس بوسائل التكنولوجيا الحديثة وبين زيادة قدرة الطلبة التحصيلية والإبداعية".
٢. الفرضية البديلة Alternative Hypothesis : H_1 : فهي العلاقة السلبية (بالنفي) بين المتغيرين مثل:
"توجد علاقة سلبية بين إفتقار المدارس لوسائل التكنولوجيا الحديثة وبين انخفاض قدرة الطلبة التحصيلية والإبداعية".

مكونات الفرضية:

تتكون الفرضية عادة من متغيرين (Variable) أساسيين هما:
المتغير المستقل Independent variable : هو المنطلق لفرضيات البحث والذي يؤثر بطريقة إيجابية أو سلبية في المتغير التابع. ولا يتواجد المتغير المستقل (المؤثر) إلا بوجود المتغير التابع (المتأثر)، وأنه أية زيادة في المتغير المستقل يقابلها زيادة أو نقصان في المتغير التابع.
المتغير التابع Dependent variable: يعتبر الأساس في اهتمام الباحث وفي تحديد مشكلة البحث، حيث يحاول الباحث بذل جهد واف لفهم هذا النوع من المتغيرات وطبيعتها. ويعتبر كحقيقة قابلة للتطبيق والاختبار وعلى هذا الأساس فإن تحليل المتغير التابع والتحري عنه وإيجاد ماهية المتغيرات التي تؤثر فيه هو أساس البحث، ومن الأمثلة على هذه المتغيرات:

- التحصيل الدراسي في المدارس الثانوية يتأثر بشكل كبير بالتدريس الخصوصي خارج المدرسة. المتغير المستقل في الفرضية هو "التدريس الخصوصي" والمتغير التابع هو التحصيل الدراسي المتأثر بالتدريس

الخصوصي، والذي يحصل كنتيجة له. فإذا طبقنا الفرضية الصفرية التي تعني العلاقة السلبية، على المثال فسيكون "لا توجد علاقة بين التدريس الخصوصي والتحصيل الدراسي".

- التفكك العائلي المتمثل بالطلاق (متغير مستقل)، وأثره على ارتفاع نسبة الجريمة عند المراهقين (متغير تابع).
- قلة الحوادث المرورية في الشوارع (متغير مستقل)، وزيادة عدد أفراد شرطة السير (متغير تابع).

فوائد الفرضيات وأهميتها:

هنالك عدد من الفوائد يمكن تحديدها بالآتي:

١. تساعد الفرضيات في تحديد أبعاد المشكلة أمام الباحث تحديداً دقيقاً يمكنه من دراستها وتناولها بعمق.

٢. تمثل الفرضيات القاعدة الأساسية لموضوع البحث والتي تجعل من السهل اختيار الحقائق المهمة واللازمة لحل المشكلة، وعدم التخبط والمتاهة وجمع كميات من المعلومات الفائضة عن الحاجة دون هدف.

٣. تقود الفرضيات الباحث إلى توجيه عملية التحليل والتفسير العلمي على أساس أن العلاقات المفترضة بين المتغيرات المستقلة والتابعة تدل الباحث إلى ما يجب أن يقوم به.

٤. تمكن الفرضيات الباحث من استنباط النتائج، حيث انه سيصل إلى الاستنتاج الذي يؤكد له بأن الفرض صحيح أو غير صحيح.

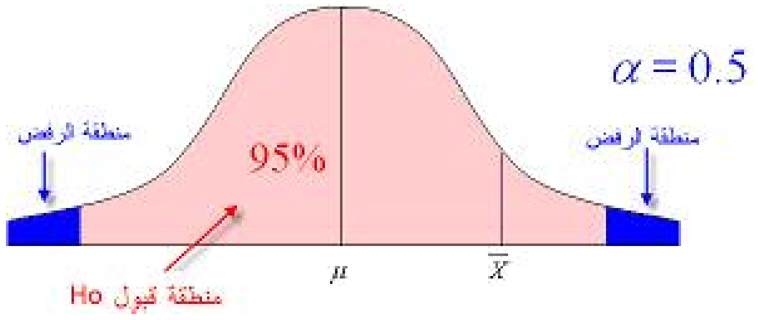
٥. تؤدي الفرضية إلى توسيع المعرفة، باعتبارها أداة فكرية يستطيع الباحث عن طريقها الحصول على حقائق تحفز باحثين آخرين إلى المزيد من البحوث الجديدة.

٦. أي دراسة (بحث، أو رسالة ماجستير، أو أطروحة دكتوراه) لن تأخذ مداها وشكلها البحثي العلمي إلا إذا ارتبطت بتحديد الباحث للمتغيرات التابعة و المستقلة.

يعتبر اختبار الفرضيات من أهم فروع الإحصاء، من اجل اختبار صحة الفرضية من بطلانها يتم اختبار الفرضيات بعدة طرق ومنها :-

- اختبار T
- اختبار Z

- هدف الاختبار هو:رفض الفرضية العدمية H0 إذا كانت خاطئة أو قبولها إذا كانت صحية، ولكن قد ينشئ أخطاء مثل:
- رفض الفرضية العدمية بينما هي صحيحه بالخطأ من النوع الأول (Type 1 error) أو خطأ (α).
- قبول الفرضية العدمية بينما هي خاطئة وتسمى هذه الحالة بالخطأ من النوع الثاني (Type 2 error) أو خطأ (β).



☑ ٣,٥ الارتباط Correlation

تستخدم مقاييس التشتت والنزعة المركزية من اجل إجراء حسابات على متغير واحد من نفس النوع فمثلاً كنا نحسب الوسط الحسابي والانحراف المعياري وغيرها لأطوال مجموعة من الطلاب أو لأوزانهم ولكن كلاً على حده. أما في الارتباط فهو دراسة متغيرين ومعرفة العلاقة بين هذين المتغيرين. كأن نأخذ مثلاً معدل مجموعة من الطلاب في الثانوية و معدلاتهم التراكمية ونجد العلاقة بين المعدلين. وهنا يلزمنا عمودين من البيانات أحدهما يمثل معدل الثانوية والآخر يمثل المعدل التراكمي، فإذا كان المتغير x يمثل معدل الثانوية والمتغير y يمثل معدل الطالب التراكمي فإننا تستطيع التعبير عن البيانات على شكل أزواج مرتبة (x,y) حيث الزوج (x_1,y_1) يمثل معدل الثانوية والتراكمي لطالب الأول وهكذا.

الهدف من دراسة وتحليل الارتباط هو معرفة العلاقة بين المتغير x والمتغير y ومعرفة نوع هذه العلاقة كأن تكون

- طردية (موجبة) (بزيادة قيمة أحد المتغيرين تزداد قيمة الآخر).
- عكسية (سالبة) (بزيادة أحدهما يقل الآخر).

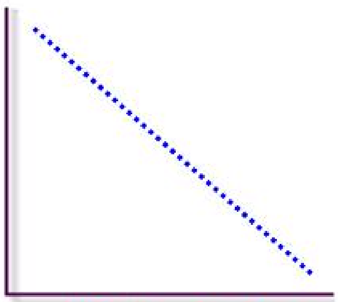
تستطيع معرفة نوع هذه العلاقة بين المتغيرين من خلال الرسم البياني (مستوى ديكارتي) حيث يمثل المتغير x محور السينات والمتغير y محور الصادات، ولقياس لقوة العلاقة الخطية بين المتغيرين ونوعها نستخدم معامل الارتباط الخطي Coefficient of Linear Correlation (r) عندما يكون معامل الارتباط r موجباً تكون العلاقة بين المتغيرين طردية وعندما يكون سالباً تكون العلاقة عكسية، أما عندما تكون قيمته صفراً فنقول أنه لا يوجد ارتباط بين المتغيرين وهو نوعان:

- معامل ارتباط بيرسون الخطي (Pearson's Correlation Coefficient).
- معامل ارتباط سبيرمان للرتب (Spearman's coefficient of rank correlation).

تتراوح قيمة معامل الارتباط r بين -1 و 1 فعندما تكون:

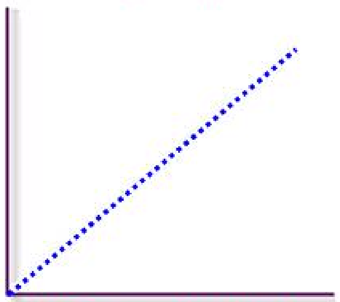
- $r = 1$ فإن الارتباط بين المتغيرين يكون خطي، موجب كامل (perfect positive linear) أي أن العلاقة طردية قوية.
- $r = -1$ فإن الارتباط يكون خطي، سالب كامل (perfect negative(inverse) linear) العلاقة عكسية قوية.
- $r = 0$ فإن ذلك يعني عدم وجود ارتباط خطي بين المتغيرين وكامل اتجهنا من ١- إلى ١+ تزداد قوة العلاقة.
- وجود ارتباط خطي كامل أو غير كامل بين متغيرين لا يعني السببية (causation) أي أن X سبباً في حدوث Y ولكن يعني أن هناك علاقة خطية بين هذين المتغيرين (بزيادة أحدهما يزداد الآخر أو ينقص).

r= -1



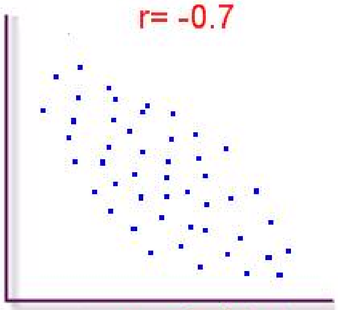
ارتباط خطي سالب كامل
(عكسي قوي)

r = +1



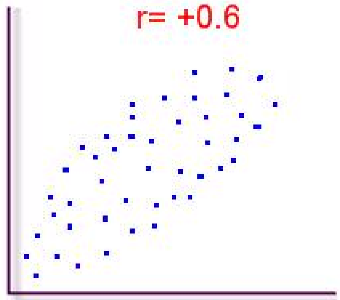
ارتباط خطي موجب كامل
(طردي قوي)

r= -0.7



ارتباط عكسي قوي

r= +0.6



ارتباط طردي قوي

➢ ٣,٥,١ معامل ارتباط بيرسون الخطي (Pearson's Correlation Coefficient).

يعطى بالمعادلة:

$$r_{xy} = r_{yx} = \frac{\sum_{1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)\ S_x\ S_y}$$

حيث : n : عدد المشاهدات في كل من X و Y

$\bar{x}$: الوسط الحسابي (Mean) لقيم X

$\bar{y}$: الوسط الحسابي (Mean) لقيم Y

S_x: الانحراف المعياري(Standard Deviation) لقيم X،

S_y: الانحراف المعياري (Standard Deviation) لقيم Y

$$S_x = \sqrt{\frac{\sum_{1}^{n}(x_i - \bar{x})^2}{n-1}}$$

وهناك صيغ أخرى لهذا القانون اسهل في عمل الحسابات ومنها :

$$r_{xy} = r_{yx} = \frac{\sum_{1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{1}^{n}(x_i - \bar{x})^2 \sum_{1}^{n}(y_i - \bar{y})^2}}$$

مثال ٢٠: احسب معامل الارتباط بين معدلات طلاب في الثانوية و معدلاتهم التراكمية المبينة في الجدول ٣,١١ ، ثم بين هل هناك علاقة وارتباط بين المتغيرين وما هو نوعها؟

جدول ٣,١١

معدل الثانوية(X)	المعدل التراكمي(Y)
٧٣	٧٥
٨٢	٧٩
٧٧	٧٥
٨٨	٨٨
٨٠	٨١
٧٩	٧٠
٧٥	٦٥
٧٦	٨٤
٨٨	٩٠
٩٠	٨٠

الحل :

نجد أولاً الوسط الحسابي لكل من X و Y ويساوي 80.8 و 78.7 على التوالي. ثم بعد ذلك نملأ الجدول أدناه ونطبق القانون الخاص بالارتباط:

جدول ٣,١٢

NO.	X	Y	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X})(Y - \bar{Y})$	$(X - \bar{X})^2$	$(Y - \bar{Y})^2$
١	٧٥	٧٣	-7.8	-3.7	28.86	60.84	13.69
٢	٧٩	٨٢	1.2	0.3	0.36	1.44	0.09
٣	٧٥	٧٧	-3.8	-3.7	14.06	14.44	13.69
٤	٨٨	٨٨	7.2	9.3	66.96	51.84	86.49
٥	٨١	٨٠	-0.8	2.3	-1.84	0.64	5.29
٦	٧٠	٧٩	-1.8	-8.7	15.66	3.24	75.69
٧	٦٥	٧٥	-5.8	-13.7	79.46	33.64	187.69
٨	٨٤	٧٦	-4.8	5.3	-25.44	23.04	28.09
٩	٩٠	٨٨	7.2	11.3	81.36	51.84	127.69
١٠	٨٠	٩٠	9.2	1.3	11.96	84.64	1.69
SUM	808	787	0	0	271.4	325.6	540.1

$$r_{xy} = r_{yx} = \frac{\sum_{1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{1}^{n}(x_i - \bar{x})^2 \sum_{1}^{n}(y_i - \bar{y})^2}}$$

$$= \frac{271.4}{\sqrt{325.6 \times 540.1}}$$

$$= \frac{271.4}{\sqrt{175856.6}}$$

$$= \frac{271.4}{419.4} = 0.65$$

نلاحظ أن معامل الارتباط بين المتغير الذي يمثل معدل الثانوية والمتغير الذي يمثل المعدل التراكمي هو ٠,٦٥ وهو موجب والعلاقة بينهما طردية وقوية (الارتباط خطي وموجب)، أي أن هناك علاقة قوية بين معدل الثانوية والمعدل التراكمي.

➢ ٣,٥,٢ معامل ارتباط سبيرمان للرتب (Spearman's coefficient of rank correlation):

نستخدم هذا النوع من عوامل الارتباط عندما تكون المتغيرات المقاسة غير رقمية و إنما تحمل ترتيب معين، فمثلاً لقياس تفاعل الطلبة مع المدرس يمكن تصنيفهم وفق ترتيب معين فنقول هذا الأول وذلك الثاني وهكذا. وأكثر ما يستخدم هذا النوع من الارتباط في المسائل المتعلقة بالاقتصاد والإدارة والتربية.

و يقاس حسب العلاقة التالية :

$$r_S = 1 - \frac{6\sum_{1}^{n} d_i^2}{n(n^2 - 1)}$$

حيث : n : عدد المشاهدات (البيانات) في كل متغير .

d : الفرق بين رتب المتغيرين x و y.

- يصلح هذا القانون إذا كان عدد أزواج البيانات أقل من 30 (عينة أقل من 30)
- نقوم بداية بترتيب البيانات من نوع x بحيث نعطي الرتبة 1 مثلاً لأعلى قيمة وهكذا، وكذلك نفعل نفس الشيء لقيم المتغير y حيث نقوم بترتيبها وإعطائها رتب حسب قيمها.

• إذا تكررت نفس القيم عدة مرات فإننا نقوم بترتيب البيانات كما لو أنه لا يوجد فيها تكرار ثم نحسب الوسط الحسابي لرتب كل مجموعة من البيانات المتساوية ونعتبر الوسط الحسابي هو رتبة كل من البيانات المتساوية في هذه المجموعة.

مثال ٢١: الجدول التالي يبين تقدير سبعة طلاب في مبحثي اللغة العربية واللغة الإنجليزية، احسب معامل الارتباط للرتب (سبيرمان) بينهما.

جدول ٣,١٣

رقم الطالب	١	٢	٣	٤	٥	٦	٧	٨
التقدير في اللغة العربية	جيد	جيد جدا	ممتاز	جيد	متوسط	ضعيف	ممتاز	جيد
التقدير في اللغة الإنجليزية	جيد جدا	ممتاز	جيد جدا	متوسط	متوسط	جيد	ممتاز	متوسط

الحل : نملأ بداية الجدول بالبيانات المطلوبة وحسب التالي :

جدول ٣,١٤

رقم الطالب	تقدير اللغة العربية	تقدير اللغة الإنجليزية	الرتبة المبدئية للعربية	الرتبة المبدئية للإنجليزي	رتبة اللغة العربية	رتبة اللغة الإنجليزية	فرق الرتب(d)	d^2
١	جيد	جيد جدا	٣	٣	٤	٣,٥	0.5	0.25
٢	جيد جدا	ممتاز	٢	١	٢	١,٥	0.5	0.25
٣	ممتاز	جيد جدا	١	٤	١	٣,٥	-2.5	6.25
٤	جيد	متوسط	٤	٦	٤	٦,٥	-2.5	6.25
٥	متوسط	متوسط	٦	٧	٦	٦,٥	-0.5	0.25
٦	ضعيف	جيد	٧	٥	٧	٥	2	4
٧	جيد	ممتاز	٥	٢	٤	١,٥	2.5	6.25
المجموع							٠	23.5

وبتطبيق معامل الارتباط للرتب على البيانات الموجودة في الجدول أعلاه نجد أن :

$$r_S = 1 - \frac{6\sum_{1}^{n} d_i^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 23.5}{7(7^2-1)}$$

$$= 1 - \frac{141}{336}$$

$$= 1 - .42$$

$$= 0.58$$

نلاحظ من الناتج أن معامل الارتباط للرتب للمتغيرين موجب تساوي (0.58)، أي أن هناك علاقة خطية موجبة (طردية) متوسطة بين المتغيرين.

☑ ٣,٦ الانحدار البسيط Linear Regression

هي المعادلة التي تربط بين المتغيرين في حال وجود علاقة وارتباط بينهما بحيث إذا عرفنا قيمة المتغير الأول تستطيع أن نتنبأ بقيمة المتغير الثاني، كالتنبؤ بمعدل طالب في الجامعة، إذا عرفت معدله في الثانوية العامة، تكتب معادلة الانحدار على الشكل $Y = \alpha + \beta X$ حيث:

Y المتغير التابع الذي نريد التنبؤ بقيمته و

X المتغير المستقل والذي نعرف قيمته. α، β ثوابت

$$\beta = \frac{\sum_{1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{1}^{n} x_i^2 - n\bar{x}^2}$$

$$\alpha = \bar{y} - \beta\bar{x}$$

مثال ٢٢: أوجد معادلة الانحدار بين معدل الطلبة في الثانوية العامة ومعدل الجامعة من البيانات الواردة في مثال (٢٠)، وما قيمة المعدل التراكمي لطالب حصل على معدل بالثانوية ٧٢ ؟.

الحل: عرفنا أن الوسط الحسابي لكل من X و Y ويساوي 80.8 و 78.7 على التوالي، وبتطبيق معادلة الانحدار على البيانات الموجودة في الجدول نجد أن:

جدول ٣,١٥

NO.	X	Y	X^2	XY
١	٧٣	٧٥	5329	5475
٢	٨٢	٧٩	6724	6478
٣	٧٧	٧٥	5929	5775
٤	٨٨	٨٨	7744	7744
٥	٨٠	٨١	6400	6480
٦	٧٩	٧٠	6241	5530
٧	٧٥	٦٥	5625	4875
٨	٧٦	٨٤	5776	6384
٩	٨٨	٩٠	7744	7920
١٠	٩٠	٨٠	8100	7200
SUM			65612	63861

$$\beta = \frac{63861 - 10\times 80.8\times 78.7}{65612 - 10\times(80.8)^2} = \frac{271.4}{325.6} = 0.83$$

$$\alpha = 78.7 - 0.83(80.8) = 11.64$$

$$\therefore Y = 11.64 + 0.83X$$

$$Y = 11.64 + 0.83(72) = 71.4$$

نعوض الآن بمعادلة الانحدار من اجل التنبؤ بقيمة المعدل التراكمي للطالب:

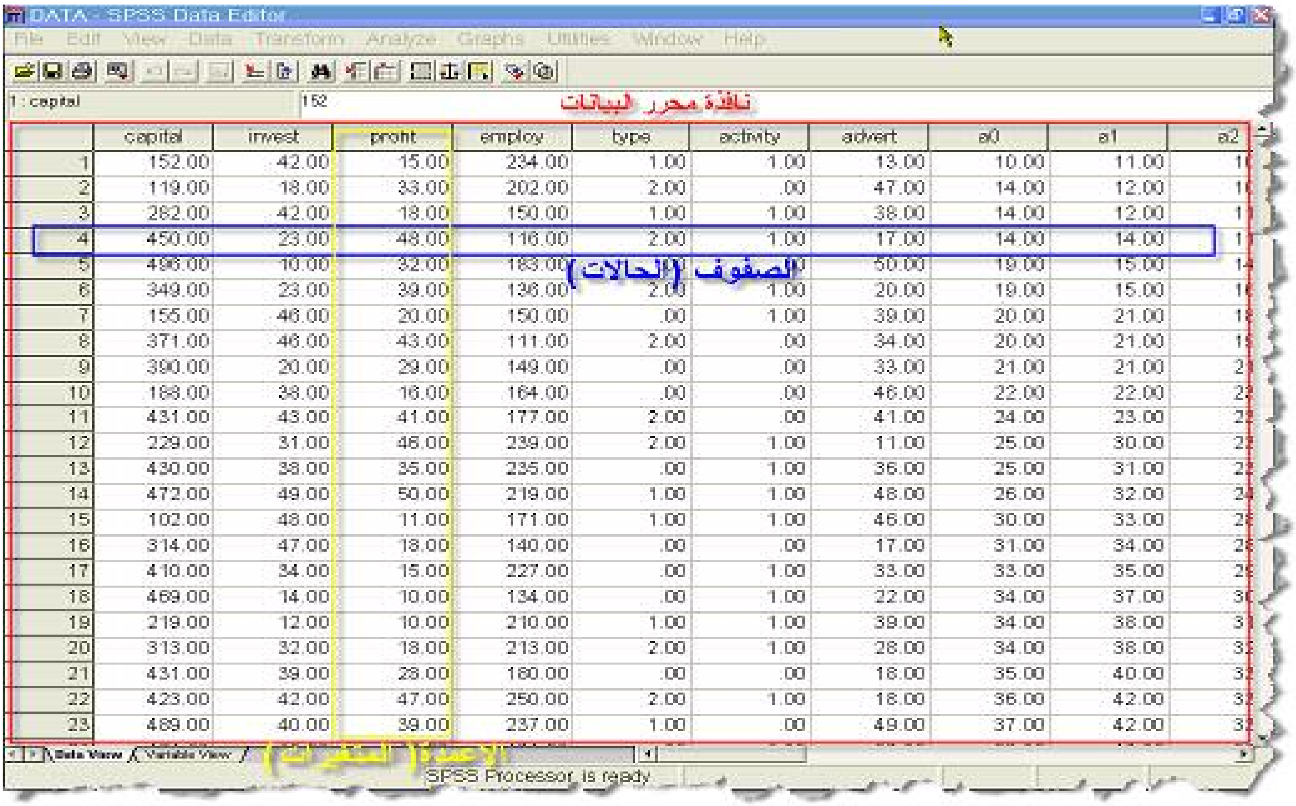
- **ننهي الجزء المتعلق بمبادئ الإحصاء وننتقل إلى الجزء العملي المتعلق ببرنامج SPSS.**

☑ ٣,٧ نوافذ برنامج SPSS:

يقدم برنامج SPSS عدة أنواع من النوافذ أهمها:

➢ ٣,٧,١ نافذة محرر البيانات Data Editor Window:

وهذه الشاشة شبيها إلى حد كبير بشاشة العمل Spreadsheet في برنامج الجداول الإلكترونية Excel، ومن خلال هذه الشاشة تستطيع إدخال وتحرير البيانات إلى البرنامج لتحليلها إحصائيا فيما بعدد، وتفتح هذه النافذة تلقائيا عند تشغيل البرنامج، وتخزن البيانات في ملف ذو امتداد (*.sav) ، كما في الشكل (١-٣).



	capital	invest	profit	employ	type	activity	advert	a0	a1	a2
1	152.00	42.00	15.00	234.00	1.00	1.00	13.00	10.00	11.00	[illegible]
2	119.00	18.00	33.00	202.00	2.00	.00	47.00	14.00	12.00	[illegible]
3	282.00	42.00	18.00	150.00	1.00	1.00	38.00	14.00	12.00	[illegible]
4	450.00	23.00	48.00	116.00	2.00	1.00	17.00	14.00	14.00	[illegible]
5	496.00	10.00	32.00	183.00	[illegible]	[illegible]	50.00	19.00	15.00	[illegible]
6	349.00	23.00	39.00	136.00	2.00	1.00	20.00	19.00	15.00	[illegible]
7	155.00	46.00	20.00	150.00	.00	1.00	39.00	20.00	21.00	[illegible]
8	371.00	46.00	43.00	111.00	2.00	.00	34.00	20.00	21.00	[illegible]
9	390.00	20.00	29.00	149.00	.00	.00	33.00	21.00	21.00	[illegible]
10	188.00	38.00	16.00	164.00	.00	.00	46.00	22.00	22.00	[illegible]
11	431.00	43.00	41.00	177.00	2.00	.00	41.00	24.00	23.00	[illegible]
12	229.00	31.00	46.00	239.00	2.00	1.00	11.00	25.00	30.00	[illegible]
13	430.00	38.00	35.00	235.00	.00	1.00	36.00	25.00	31.00	[illegible]
14	472.00	49.00	50.00	219.00	1.00	1.00	48.00	26.00	32.00	[illegible]
15	102.00	48.00	11.00	171.00	1.00	1.00	46.00	30.00	33.00	[illegible]
16	314.00	47.00	18.00	140.00	.00	.00	17.00	31.00	34.00	[illegible]
17	410.00	34.00	15.00	227.00	.00	1.00	33.00	33.00	35.00	[illegible]
18	469.00	14.00	10.00	134.00	.00	1.00	22.00	34.00	37.00	[illegible]
19	219.00	12.00	10.00	210.00	1.00	1.00	39.00	34.00	38.00	[illegible]
20	313.00	32.00	18.00	213.00	2.00	1.00	28.00	34.00	38.00	[illegible]
21	431.00	39.00	28.00	180.00	.00	.00	18.00	35.00	40.00	[illegible]
22	423.00	42.00	47.00	250.00	2.00	1.00	18.00	36.00	42.00	[illegible]
23	489.00	40.00	39.00	237.00	1.00	.00	49.00	37.00	42.00	[illegible]

الشكل (١-٣) نافذة تحرير البيانات

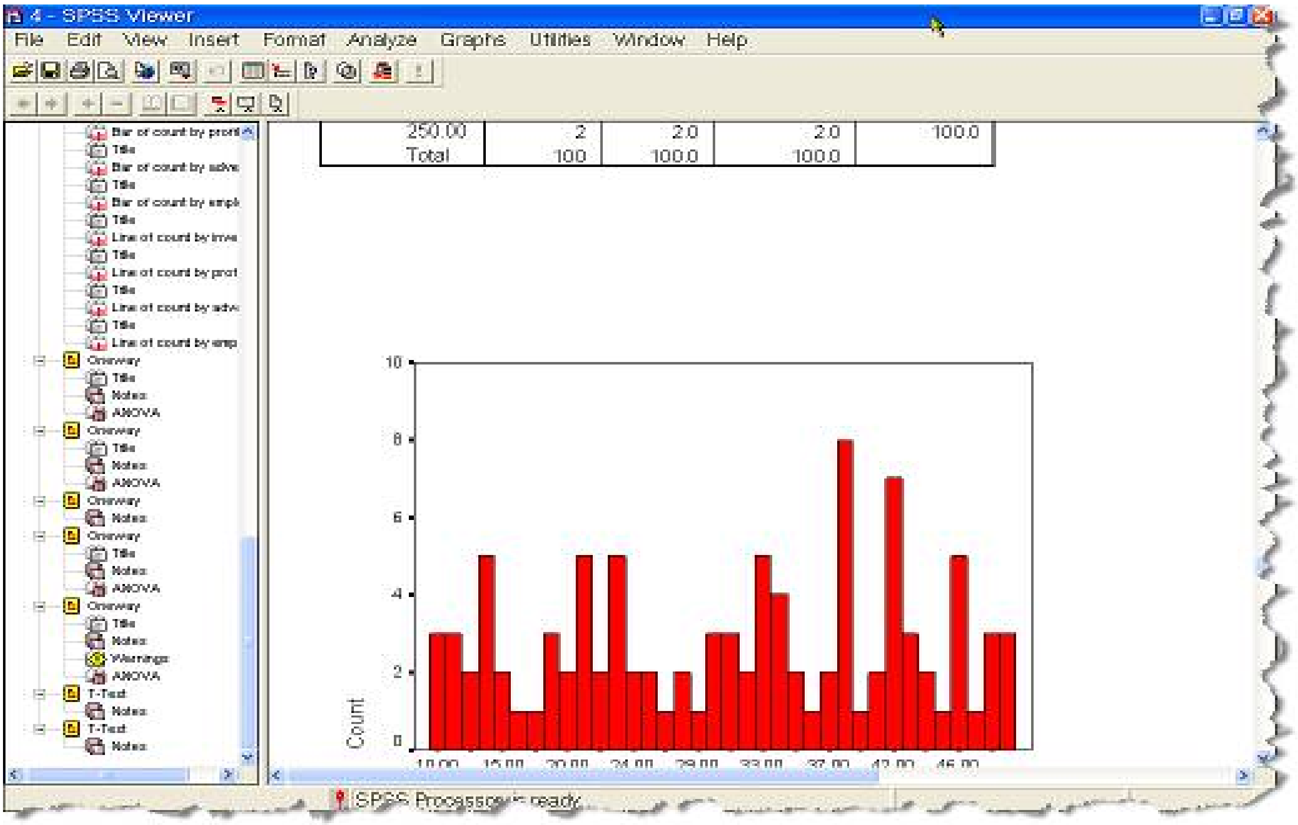
تتكون النافذة في الشكل السابق من:

الأعمدة: وهي المتغيرات الموجودة لديك ، كاسم الشركة، أو رأس المال وغيرها، وأسماء هذه الصفوف هي أرقام يمثل أخرها عدد أفراد العينة، وتكون خافتة عند إنشاء ملف جديد.

الصفوف (الحالات): وهي أفراد العينة الموجودة لديك، وهي قيمة معينة تقابل أسماء المتغيرات، وتبدأ أسماء هذه الأعمدة عند إنشاء ملف جديد باسم VAR اختصارا لــ Variable أي متغير.

➢ ٣,٧,٢ نافذة المخرجات Output Window:

وفيها يتم ظهور نتائج العمليات الإحصائية والرسوم البيانية وبالإمكان إجراء التعديلات اللازمة على هذه النتائج في هذه النافذة، وتخزن النتائج في ملف ذو امتداد (*.spo) ، ومن الممكن فتح اكثر من ملف مخرجات في نفس الوقت عن طريق قائمة ملف File واختيار القائمة الفرعية New ثم إنتق الأمر Output ، كما في الشكل (٢-٣).



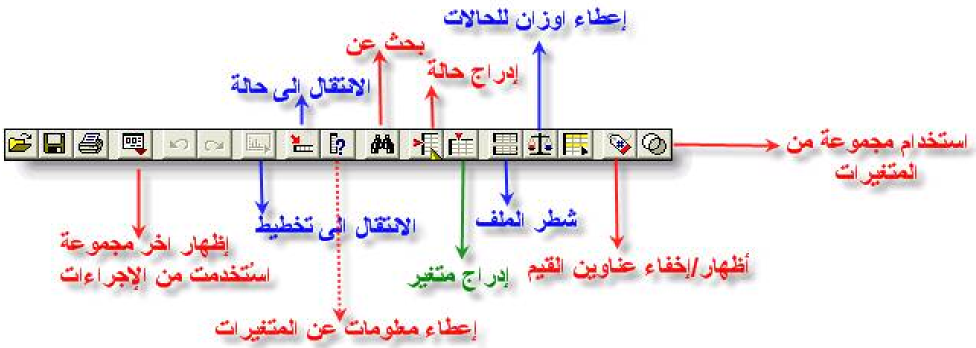
الشكل (٢-٣) نافذة المخرجات

➢ ٣,٧,٣ نافذة التعليمات Syntax Windows:

وهنا يتم كتابة تعليمات لعمليات إحصائية مختلفة تنفذ وتعدل في وقت لاحق. تخزن هذه التعليمات في ملف ذو امتداد (*.sps).

☑ ٣,٨ شريط الأدوات :

يحتوي هذا الشريط على أيقونات Icons تمثل وظائف متعددة، الشكل(٣-٣) يوضح وظيفة كل رمز في الشريط.



الشكل (٣-٣) شريط الأدوات

☑ ٣,٩ القوائم الرئيسة:

وهي أداة مختصرة للوصول إلى الأوامر في الأنظمة التي تعمل تحت نظام التشغيل النوافذ، ويوجد في SPSS بعشر قوائم رئيسة تتخللها قوائم فرعية، وهذه القوائم هي:

١) قائمة ملف File Menu: يتم من خلال هذه القائمة إنشاء ملف جديد New، فتح ملف مخزن Open من عدة أنواع، تخزين Save، طباعة الملفات Print، والخروج من النظام Exit.

٢) قائمة تحرير Edit Menu: هنا يتم النسخ، النقل، والبحث.

٣) قائمة عرض View Menu: هنا يتم إظهار شريط الأدوات Toolbar وشريط الحالة Status Bar أو إخفائهما، إخفاء خطوط الشبكة Gridlines ، تغير نوع الخط، إظهار وإخفاء عناوين القيم Value Label.

٤) قائمة بيانات Data Menu: من خلال هذه القائمة يتم تّعريف العمليات الحسابية وتغير أسمائها، وعمليات الفرز، والكثير من العمليات الأخرى.

٥) قائمة التحويلات Transform Menu: في هذه القائمة يتم القيام بالعمليات الحسابية، كاستخدام الدوال الإحصائية.

٦) قائمة الإجراءات الإحصائية Statistics Menu: تحتوي هذه القائمة على أدوات كثيرة تستخدم في عمليات التحليل الإحصائي، كحساب الانحرافات المعيارية ومعادلات الانحدار والارتباط وغيرها الكثير.

٧) قائمة الرسومات Graphs Menu: هنا يتم إنشاء الرسوم البيانية بأشكالها المختلفة.

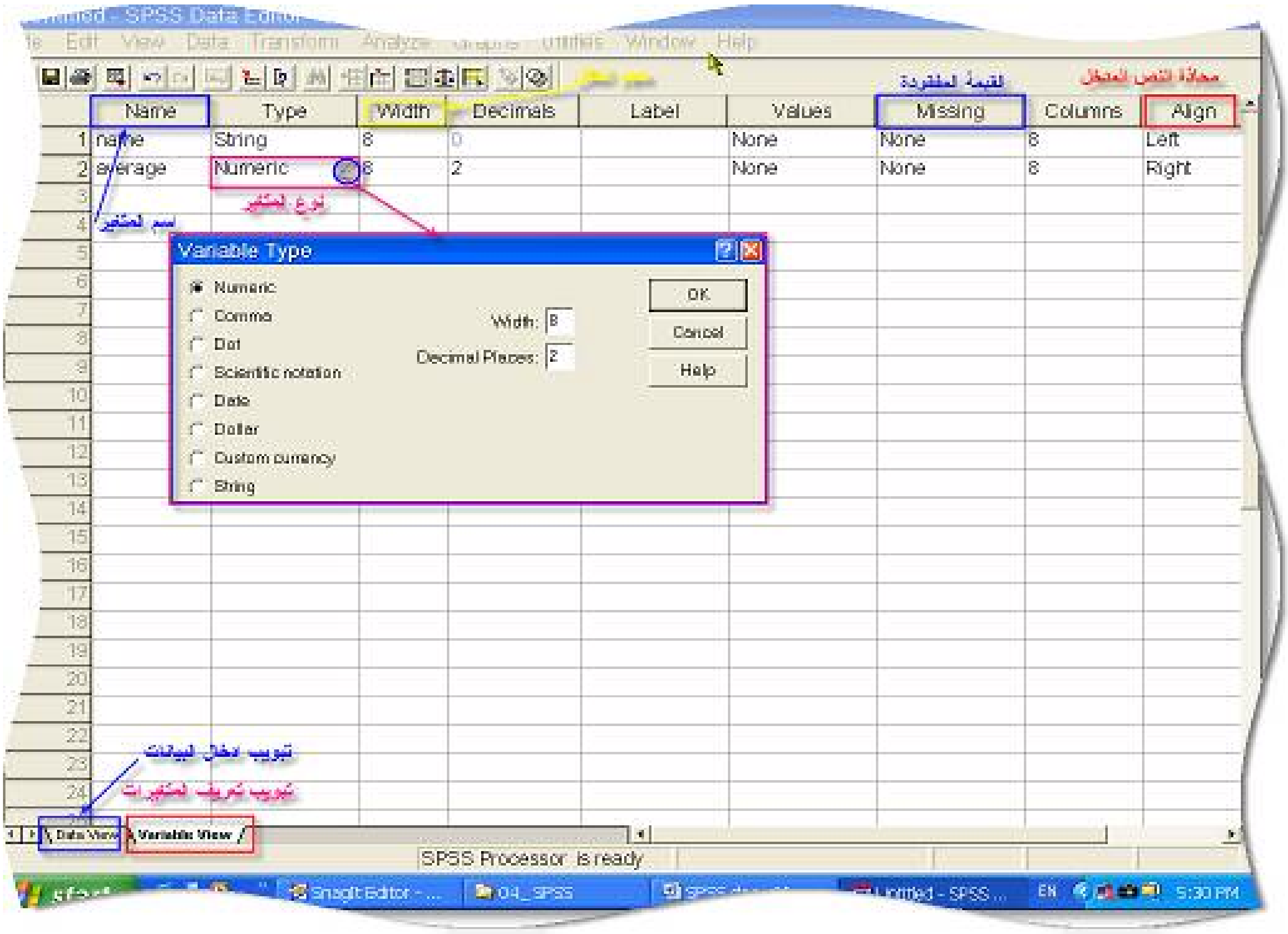
٨) قائمة الأدوات Utilities Menu: تحتوي هذه القائمة على معلومات دقيقة عن الملف المستخدم والمتغيرات والمجموعات Sets المتعلقة بهذه المتغيرات.

٩) قائمة إطار Window Menu: هنا يمكن التنقل بين النوافذ المفتوحة.

١٠)قائمة المساعدة Help Menu: هنا يمكن الحصول على المعلومات المفيدة كحل المشاكل المتعلقة بنظام SPSS.

تعريف المتغيرات:

قبل عملية إدخال البيانات يجب تعريف المتغيرات، لتعريف المتغيرات ننتقل إلى التبويب Variable View الموجود في اســفل الطرف الأيسر من نافذة تحرير البيانات، كما في الشكل(٤-٣). يجب تعريف المتغير عن طريق إعطاء قيم للمعلومات القاموسية Dictionary Information الخاصة بالمتغير ومن هذه المعلومات:



الشكل (٤-٣) نافذة تعريف المتغيرات

- الاسم Name: أسماء هذه المتغيرات، لا تحوي على رموز خاصة مثل @، $...الخ، ولا يزيد حجمها عــن ثمانيــة حروف.
- النوع Type: عند النقر على المربع المنقط في العمود Type يظهر مربع الحوار المبين بالشكل(٤-٣) والذي يحتوي على عدة أنواع من البيانات هي:
  - Numeric: عدد حقيقي أو صحيح.
  - Comma: عدد يحتوي على فاصلة (،) بعد كل ٣ خانات مثل ١٢٤،٥٤١،١٠١.
  - Dot: عدد حقيقي يتم الفصل بين جزئية الحقيقي والصحيح بالفاصلة (،) مثل ١٢٤,٣٢١،٢١٢.
  - Scientific Notation: عدد كبير جدا أو صغير جداً، مثل $3E+9 = 3*10^9$.
  - Date: متغير من نوع تاريخ أو وقت.
  - Dollar: متغير من نوع عملة الدولار.
  - Custom Currency: متغير من نوع عملة يعرفها المستخدم كاليورو أو الدينار أو غيرهما.
  - String: متغير من نوع نص.

- عرض Width: يمثل طول المتغير، وفي الأعداد تدخل الفاصلة والكسر من مجمل الحجم، وسيقرب العدد إذا كان أطول من الطول المحجوز.
- Decimal: يمثل عدد خانات الجزء الحقيقي من العدد.
- تنسيق العمود Column : يمكن تحديد عرض العمود المطلوب بزيادة أو إنقاص القيمة الموجودة في مربع column.

مثال ٢٣: أنشئ ملف جديد من نوع Data وأدخل فيه البيانات المبينة في الجدول التالي، ثم قم بتخزينها في ملف اسمه Student.

الجدول (١-٣)

رقم الطالب	العمر	الجنس	المستوى	المعدل التراكمي
٢٠٠٥١٢	١٩	ذكر	الأولى	٨٨
٢٠٠٥١٥	١٩	أنثى	الثانية	٨٢,٤
٢٠٠٤٨٨	٢٠	أنثى	الثانية	٧٨,٢
٢٠٠١٨٩	٢٢	ذكر	الرابعة	٧٢

الحل:

١) من قائمة ملف File إنتق جديد New ثم إنتق الأمـــــــــــــــــــــر Data.

الشكل (٥-٣)

في تبويب Variable View عرّف المتغيرات المستخدمة وحدد جميع قيم الحقل الخاصة بكل متغير، كما في الشكل (٦-٣).

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1	stu_num	Numeric	8	2		None	None	8	Right
2	age	Numeric	8	2		None	None	8	Right
3	sex	String	8	0		None	None	8	Center
4	level	String	8	0		None	None	8	Center
5	average	Numeric	8	2		None	None	8	Right

الشكل (٦-٣) المتغيرات

٢) إذهب إلى Data View وندخل البيانات الموجودة في الجدول، كما في الشكل(٧-٣).

	stu_num	age	sex	level	average
1	200512.0	19.00	male	1st	88.00
2	200515.0	19.00	female	2nd	82.40
3	200488.0	20.00	female	2nd	78.20
4	200189.0	22.00	male	4th	72.00

الشكل (٧-٣) البيانات بعد الإدخال

٣) إحفظ ما قمنا به بالذهاب إلى قائمة ملف والنقر على الأمر Save As. ثم إنتق Directory المراد التخزين فيه، ونضع في مربع "اسم الملف" اسم الملف وهو Student وكما قلنا سيتم تخزين البيانات في ملف من نوع(sav.*).

3.9.1 فتح ملف بيانات مخزن Open:

- من قائمة ملف إنتق الأمر Open. ثم إنتق من القائمة الفرعية نوع الملف المطلوب فتحة، خنا نريد فتح ملف من نوع Data ، كما في الشكل (٨-٣).

الشكل (٨-٣) فتح ملف

٣,٩,٢ حذف متغير (عمود)Delete Variable (column :

- حدد العمود المراد حذفه بالنقر عليها بالماوس ثم أضغط على Delete من لوحة المفاتيح أو إنتق Clear من قائمة Edit.

٣,٩,٣ حذف صف (حالة) Delete Row(case):

- حدد الصف المراد حذفه بالنقر عليها بالماوس ثم أضغط على Delete من لوحة المفاتيح أو إنتق Clear من قائمة Edit.

٣,٩,٤ نسخ و قص Copy & Cut:

١) حدد الصف أو العمود المراد نسخه أو نقله.

٢) من قائمة Edit إنتق Copy أو Cut.

٣) حدد المكان المراد النسخ إليه ثم إنتق Paste من نفس القائمة.

٣,٩,٥ البحث عن القيم Finding Values.

- من قائمة Edit إنتق الأمر Find، فتظهر نافذة بحث. أدخل في مربع الحوار القيمة المراد البحث عنها، ثم إنتق Find Next فينتقل المؤشر إلى تلك القيمة، أو تظهر نافذة تشير إلى عدم وجود هذه القيمة.

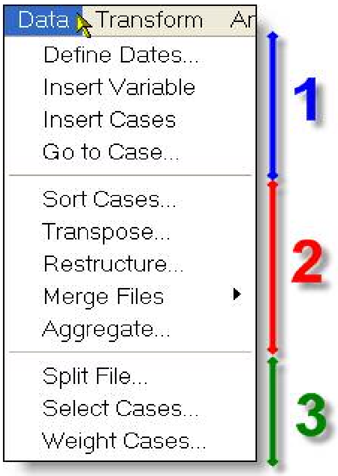
٣,١٠ قائمة البيانات Data :

تقسم قائمة البيانات كما في الشكل(٩-٣) إلى ثلاثة أقسام هي:

١) **القسم الأول:** ويحتوي على أوامر تعريف البيانات قبل أن يتم استخدامها، كإضافة صف أو عمود أو تعريف متغير.

٢) **القسم الثاني:** وهي الأوامر التي تجرى على البيانات بعد إدخالها، كترتيبها أو دمج الملفات، وتحويل الصف إلى عمود. فتظهر النتائج مباشرة.

٣) **القسم الثالث:** لا تظهر هذه الأوامر مباشر على البيانات، وإنما تظهر فيما بعد عند استخدام الأوامر الاحصائية.



الشكل (٩-٣)

➢ ٣,١٠,١ إدراج متغير (عمود) Insert Variable (column) :

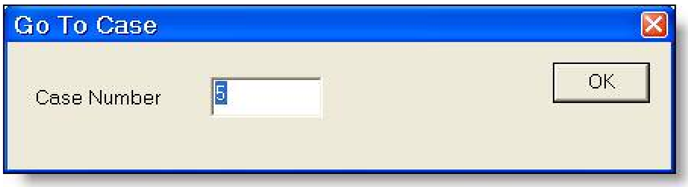
- حدد العمود المراد إدراج عمود أمامه وذلك بالنقر عليها بالماوس ثم إنتق Insert Variable من قائمة Data. أو بالنقر على رمز الإدراج في شريط الأدوات.

➢ ٣,١٠,٢ إدراج صف (حالة) Insert cases (Row):

- حدد الصف المراد إدراج صف جديد أمامه وذلك بالنقر عليها بالماوس ثم اختر Insert Cases من قائمة Data. أو بالنقر على رمز الإدراج في شريط الأدوات.

➢ ٣,١٠,٣ الذهاب إلى حالة ما Go To Case:

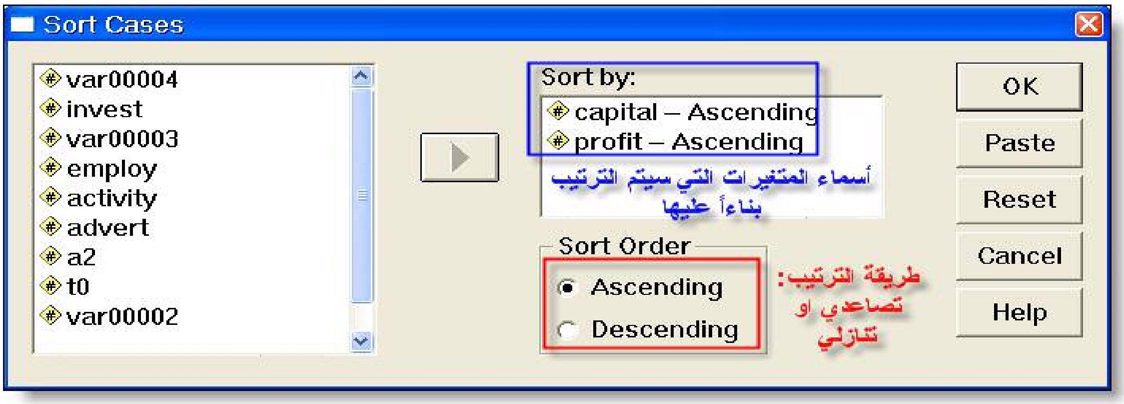
- من قائمة Data إنتق الأمر Go To Case فتظهر شاشة كما في الشكل(١٠-٣)، أدخل رقم الصف المراد الذهاب إليه ثم إنتق موافق.



الشكل (١٠-٣) نافذة الذهاب إلى

➢ ٣,١٠,٤ ترتيب البيانات Sort Cases:

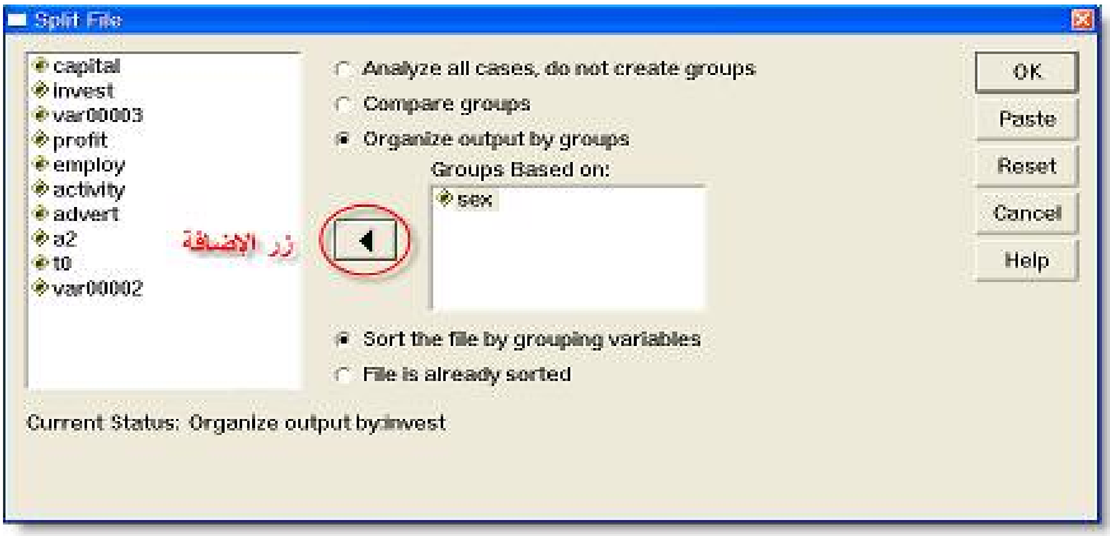
- لترتيب محتويات الصفوف، كترتيب علامات الطلبة من الأدنى إلى الأعلى. إذهب إلى قائمة Data ثم إنتق Sort Cases، ستظهر نافذة إنتق منها اسم المتغير الذي ستترتب قيم الصفوف بناءاً عليه. كما إنتق نوع الترتيب التصاعدي Ascending أو تنازلي Descending، كما في الشكل(١١-٣).



الشكل (١١-٣) نافذة الترتيب

➢ ٣,١٠,٥ تقسيم الملفات Split File.

- يوفر برنامج SPSS إمكانية تقسيم البيانات المخزنة في ملف إلى قسمين، كتقسيم البيانات حسب الجنس لدراسة كل عينة بشكل مستقل للوصول إلى نتيجة معينة. ويتم التقسيم من خلال الأمر Split File الموجود في قائمة Data فتظهر النافذة المبينة بالشكل(١٢-٣)، إنتق الخيار الثالث Organize output by groups هنا إنتق اسم المتغير الذي ستقسم البيانات على أساسه ولنفرض انه Sex نضيفه بواسطة السهم ثم أنقر على زر موافق.



الشكل (١٢-٣) نافذة تقسم البيانات

- سوف تنقسم البيانات إلى قسمين قسم ذكور وقسم إناث كما في الشكل(١٣-٣)، كما قلنا إن الأوامر الموجودة في القسم الثالث لا تظهر نتائجها مباشرة، ولكن الذي يظهر هو مجرد ترتيب حسب الجنس أما نتائج التقسيم ستظهر لاحقاً عند استخدام الأوامر الإحصائية من قائمة Statistics. فإذا قمنا بحساب انحراف المعدلات سيقوم SPSS بحساب انحرافين واحد لذكور و الأخر للإناث بشكل منفصل.

	stu_num	age	sex	level	average
1	200515	19	female	2nd	82.40
2	200488	20	female	2nd	78.20
3	200355	21	female	3rd	79.80
4	200512	19	male	1st	88.00
5	200189	20	male	4th	72.00
6	200290	22	male	4th	75.00

الشكل (١٢-٣) البيانات بعد التقسيم

➢ ٣,١٠,٦ دمج الملفات Merge Files :

١) دمج ملفان يحتويان على نفس المتغيرات مع اختلاف في الحالات (الصفوف)Add Cases:

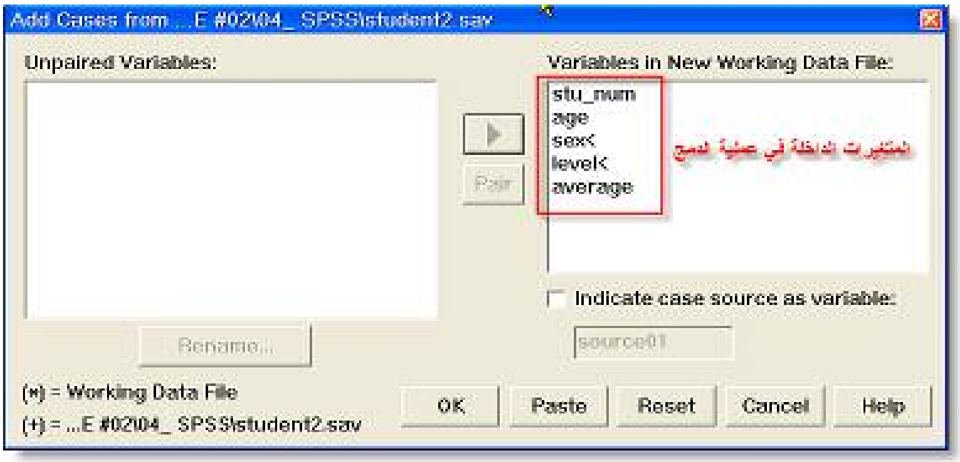
١. لدمج الملف Student1 و الملف Student2 نقوم بفتح الملف الأول Student1 ثم إذهب إلى قائمة Data ونختار منها القائمة الفرعية Merge Files ثم إنتق الأمر Add Cases.

٢. تظهر نافذة أخرى إنتق منها الملف المراد دمجه مع الملف المفتوح وهو Student2. ثم إضغط على Open.كما في الشكل(١٣-٣).

٣. تظهر نافذة جديدة، إضغط موافق، الشكل(١٤-٣).

٤. يُظهر الشكل(١٥-٣) الملف بعد الدمج، قم بتخزين هذه النتيجة بملف جديد.

الشكل (١٣-٣) نافذة إضافة ملف



الشكل (١٤-٣) نافذة اختير المتغيرات

	stu_num	age	sex	level	average
1	200452	20	female	1st	72.50
2	200419	20	male	1st	88.60
3	200333	21	female	2nd	90.80
4	200261	22	female	3rd	82.00
5	200125	23	female	3rd	73.00
6	200122	23	male	4th	77.00
7	200111	23	female	4th	80.00
8	200452	20	female	1st	72.50
9	200419	20	male	1st	88.60
10	200333	21	female	2nd	90.80
11	200261	22	female	3rd	82.00
12	200125	23	female	3rd	73.00
13	200122	23	male	4th	77.00

الشكل(١٥-٣) نتيجة الدمج

٢) دمج ملفان يحتويان على نفس الحالات (الصفوف) مع اختلاف المتغيرات Add Variables:

في هذه الحالة يتم إضافة متغيرات (أعمدة) جديدة إلى تلك المتغيرات الموجودة سابقاً، والشرط هنا أن يحتوي كلا الملفين على نفس العدد من الحالات (الصفوف). يحبذ أن تكون المتغيرات في الملفين مترابطة كأن تكون المتغيرات في الملفين لنفس الأشخاص بالترتيب نفسه، ربما يحتوي الملف الأول على اسم الطالب ورقمه الجامعي ويحتوي الملف الثاني على علاماته، وفي حالة عدم تأكدنا من هذه العلاقة نستخدم متغير يطابق بين البيانات في الملفين يدعى المتغير المفتاح Key Variable ، كأن يكون الرقم الجامعي في الملفين هو المتغير المفتاح.

	stu_num	math_101	cs_211	cs_101	ph_102
1	200452	78	80	85	75.00
2	200419	77	77	72	70.00
3	200333	85	89	85	88.00
4	200261	60	70	72	75.00
5	200125	82	80	79	78.00
6	200122	71	74	75	78.00

المتغير المفتاح

الشكل(١٦-٣) الملف sGrade

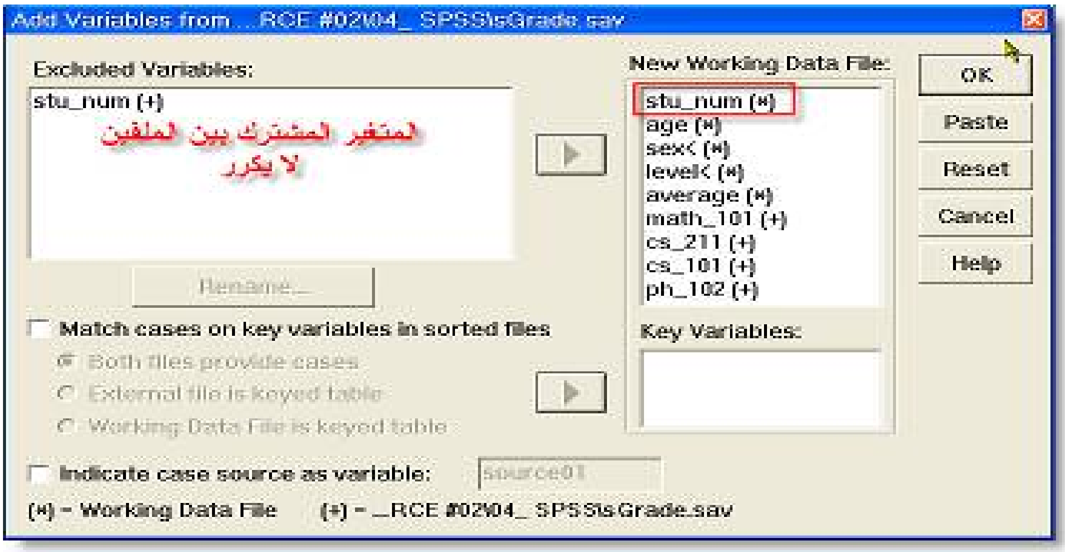
١. لدمج الملف Student1 و الملف sGrade ، نفتح الملف الأول Student1 ثم إذهب إلى قائمة Data ونختار منها القائمة الفرعية Merge Files ثم إنتق الأمر Add Variables.

٢. تظهر نافذة أخرى إنتق منها الملف المراد دمجه مع الملف المفتوح وهو sGrade. ثم إضغط على Open. الشكل(٣-١٧).

الشكل (٣-١٧) نافذة إضافة ملف

٣. تظهر نافذة جديدة، إضغط موافق، الشكل(٣-١٨).

٤. يُظهر الشكل(٣-١٩) الملف بعد الدمج، قم بتخزين هذه النتيجة بملف جديد.



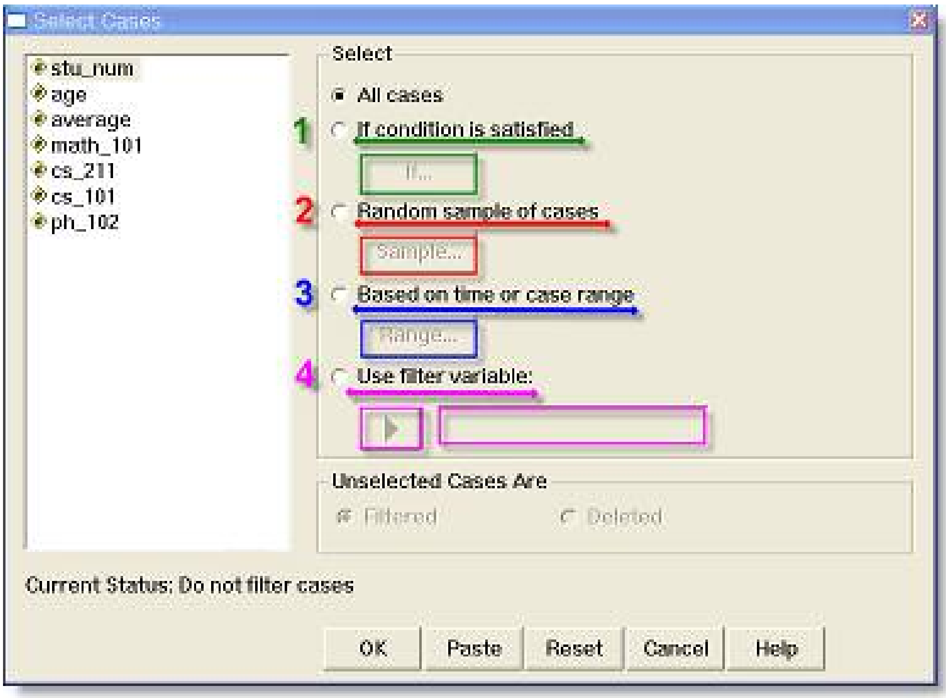
الشكل (٣-١٨) نافذة اختير المتغيرات

	stu_num	age	sex	level	average	math_101	cs_211	cs_101	ph_102
1	200512	19	male	1st	88.00	78	80	85	75.00
2	200515	19	female	2nd	82.40	77	77	72	70.00
3	200488	20	female	2nd	78.20	85	89	85	88.00
4	200355	21	female	3rd	79.80	60	70	72	75.00
5	200189	20	male	4th	72.00	82	80	79	78.00
6	200290	22	male	4th	75.00	71	74	75	78.00

الشكل(٣-١٩) نتيجة الدمج

➢ ٣,١٠,٧ اختيار الحالات Select Cases :

في بعض الأحيان يجري الباحث عملياته الإحصائية على حجم محدد من العينة (الحالات) وفق ما تتطلبه الدراسة، وعملية اختيار هذا الحجم من العينة تتم بأربعة طرق كما في الشكل (٣-٢٠) ، وهذه الطرق هي:

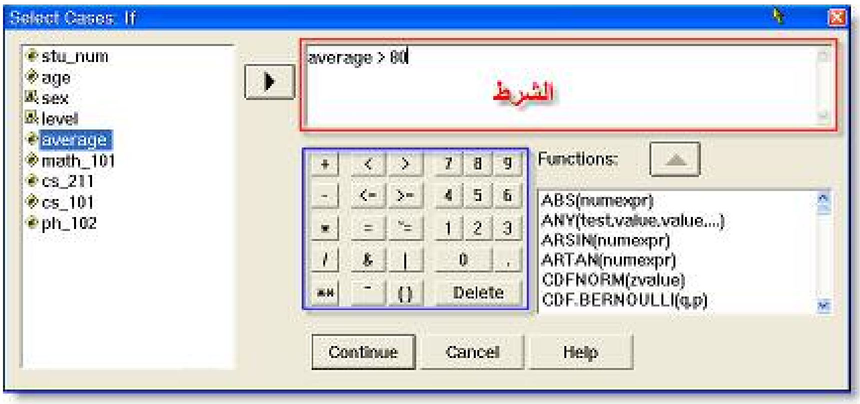


الشكل(٣-٢٠) نافذة Select Cases

١. تحقق شرط ما If condition is satisfied.

من خلال هذا الشرط تستطيع مثلاً اختيار الطلاب (الحالات) الذين يزيد معدلهم التراكمي عن ٨٠. ويتم ذلك من خلال.

- من نافذة Select Cases إنتق If condition is satisfied وننقر فوق زر If.
- تظهر لك النافذة المبينة بالشكل(٣-٢١) أكتب الشرط المطلوب ثم إضغط continue ثم Ok.



الشكل(٣-٢١) نافذة If

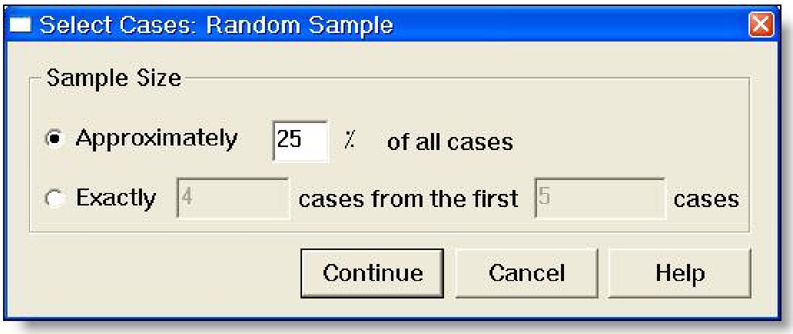
	stu_num	age	sex	level	average
1	200512	19	male	1st	88.00
2	200515	19	female	2nd	82.40
3	200488	20	female	2nd	78.20
4	200355	21	female	3rd	79.80
5	200189	20	male	4th	72.00
6	200290	22	male	4th	75.00

الشكل(٣-٢٢) ناتج استخدام طريقة الشرط

- في الشكل(٣-٢٢) نلاحظ أن SPSS قد قام بشطب الحالات التي لا تحقق الشرط السابق

٢. اختيار جزء عشوائي Random sample of cases.

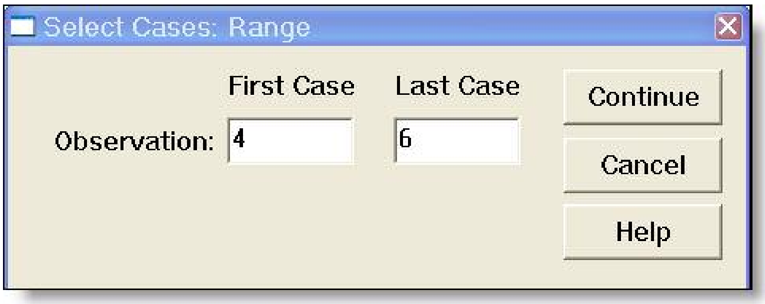
في هذه الطريقة يتم الاختيار العشوائي من العينة. أما بتحديد نسبة مئوية أو بتحديد حجم الجزء العشوائي. وذلك بالضغط على زر Sample ثم تحديد حجم الجزء. كأن إنتق من ٤ إلى ٥ حالات أو ٢٥% من حجم العينة الكلي. كما في الشكل (٢٣-٣).



الشكل(٢٣-٣) استخدام طريقة الجزء العشوائي

٣. اختيار من مدى معين Based on time or case rang.

تستطيع هنا أن تحدد فترة Rang ليتم أخيارها من العينة.وذلك بالضغط على زر Rang من نافذة Select Cases ثم تحدد هذه الفترة.



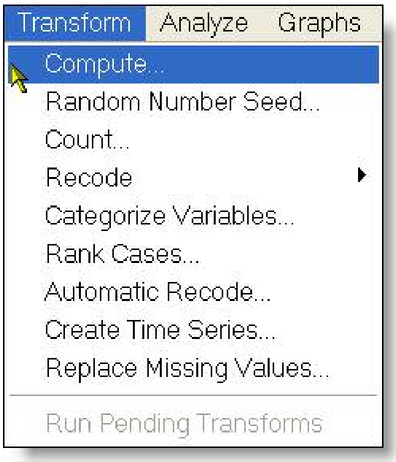
الشكل(٢٤-٣) استخدام طريقة الفترة

٤. التصفية حسب قيمة متغير Use filter variable.

في هذا الخيار تحدد المتغير المطلوب مقارنة قيمته، تُشطب الحالات التي تساوي قيمة المتغير فيها صفر، وتبقى باقي الحالات ليتم اختيارها.

☑ ٣,١١ قائمة التحويلات Transformation:

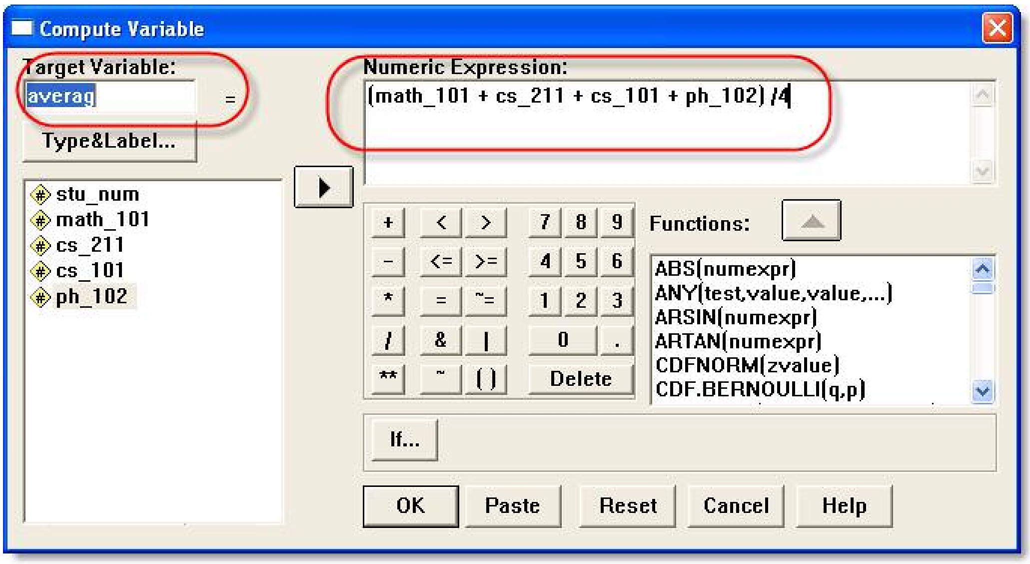
التحويل Transform هي: إنتاج متغيرات جديدة من متغيرات موجودة سابقاً. فإذا كان لديك مجموعة من علامات طلاب فكل علامة من هذه العلامات هو متغير، فإذا حسبنا الوسط الحسابي لهذه العلامات فان هذا الوسط هو تحويل تم إيجاده من متغيرات سابقه وهي العلامات، ويعتبر كذلك مجموع هذه العلامات تحويل أخر وهكذا.



الشكل(٢٥-٣) قائمة التحويلات

٣,١١,١ الأمر Compute لإجراء العمليات الحسابية:

عند النقر على هذا الأمر تظهر لك النافذة المبينة بالشكل(٢٦-٣) و التي من خلالها يمكن كتابة المعادلات المنطقية (IF) أو الدوال (الإقترانات) الرياضية Functions وكذلك يمكن استخدام الآلة الحاسبة الموجودة في هذه النافذة.



الشكل(٢٦-٣) نافذة Compute Variable

مثال ٢٤: إذا أردت أن نحسب المعدل لمجموع علامات الطلبة في ملف sGrade :

١) أدخل اسماً للعمود الجديد والذي سيمثل القيمة المحسوبة في مربع Target Variable وليكن هذا الاسم هو Average.

٢) أكتب المعادلة المناسبة في مربع Numeric Expression.

٣) ثم إضغط على زر Ok.

- سيظهر عمود جديد يحمل اسم المتغير السابق يخزن فيه قيم معدل جميع الطلبة كما في الشكل (٢٧-٣).

	stu_num	math_101	cs_211	cs_101	ph_102	averag
1	200512	78	80	85	75.00	79.50
2	200515	77	77	72	70.00	74.00
3	200488	85	89	85	88.00	86.75
4	200355	60	70	72	75.00	69.25
5	200189	82	80	79	78.00	79.75
6	200290	71	74	75	78.00	74.50

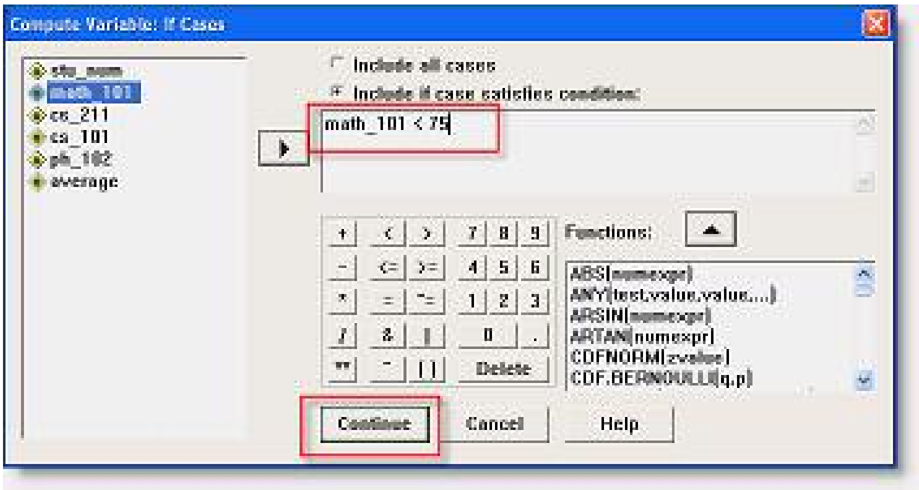
الشكل(٢٧-٣) النتيجة

مثال ٢٥ (على جملة IF): إذا أردت أن نزيد علامات الطلبة التي تقل علاماتهم في مادة Math 101 عن ٧٥ بمقدار ٥ علامات في نفس المتغير math_101 .

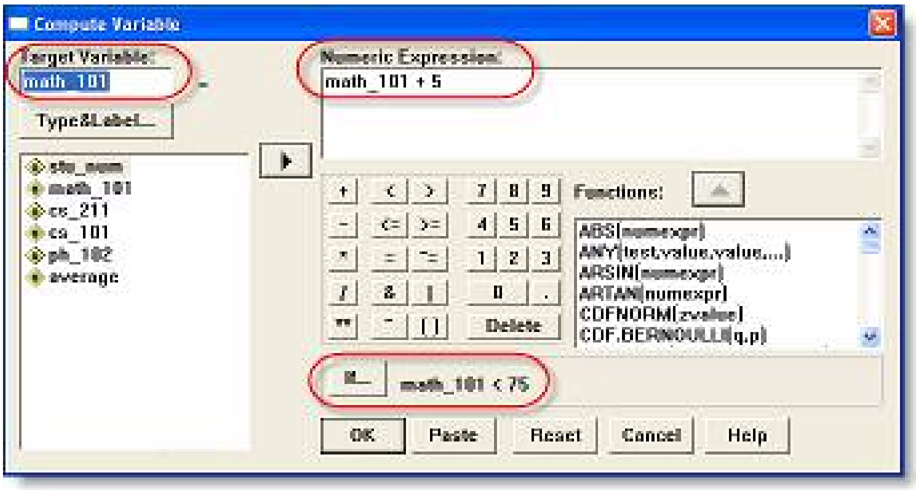
١) إضغط على زر IF في من نافذة Compute Variable الشكل(٢٦-٣).

٢) تظهر لك نافذة أخرى الشكل(٢٨-٣)، في هذه أدخل الجملة الشرطية math_101 < 75.

٣) إضغط على Continue لنعود للنافذة السابقة. الآن أدخل اسم المتغير القديم math_101 وفي المربع الأخر أدخل المعادلة المطلوبة math_101 + 5 .

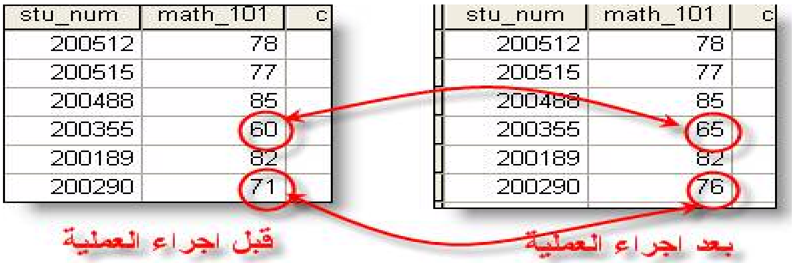


الشكل(٢٨-٣) نافذة IF



الشكل(٢٩-٣)

- ستظهر النتيجة كما في الشكل التالي (٣٠-٣) نلاحظ أن زيادة العلامات حدثت للذين قلت علاماتهم عن ٧٥:



stu_num	math_101
200512	78
200515	77
200488	85
200355	60
200189	82
200290	71

stu_num	math_101
200512	78
200515	77
200488	85
200355	65
200189	82
200290	76

الشكل(٢٩-٣) النتيجة

مثال ٢٦ (على الدوال Functions): لإيجاد المجموع الكلي لعلامات الطلاب نقوم بما يلي:

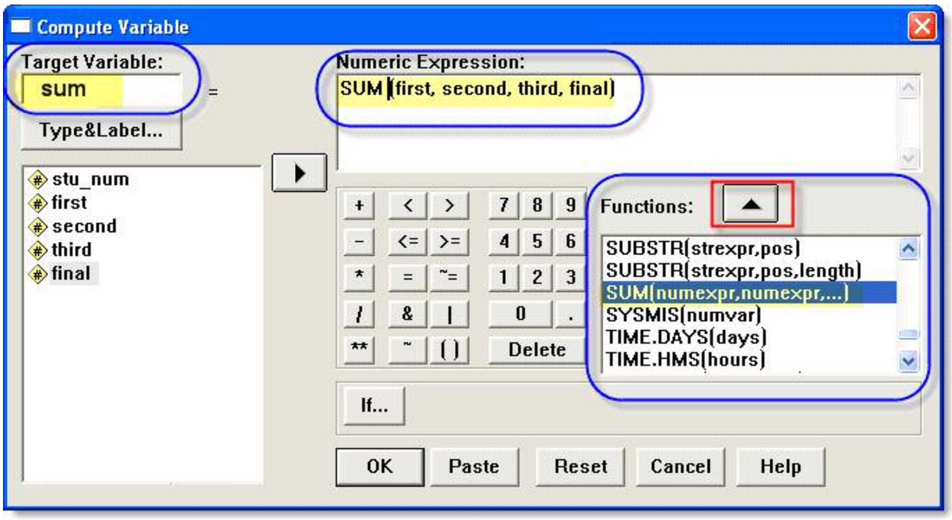
١) إنتق الدالة sum من قائمة Functions الموجودة في نافذة Compute Variable والتي تحتوي على اكثر من ٧٠ دالة كما في الشكل(٣٠-٣).

٢) أدخل إسماً للمتغير كما في السابق.

٣) إنتق الدالة SUM من مربع Functions ثم إضغط على السهم الموجود يمين كلمة Functions .

٤) أدخل القيم first ، second ، third و final داخل أقواس الدالة SUM .

٥) أنقر على Ok . فتظهر النتائج كما في الشكل(٣١-٣).



الشكل(٣٠-٣) تطبيق الدالة SUM

	stu_num	first	second	third	final	sum
1	200512	15	18	18	35	86.0
2	200515	16	17	19	38	90.0
3	200488	19	18	18	39	94.0
4	200355	14	10	12	27	63.0
5	200189	11	12	15	21	59.0
6	200290	13	17	19	25	74.0

الشكل(٣١-٣) نتيجة تطبيق الدالة SUM

➢ ٣,١١,٢ الأمر Count:

يستخدم هذا الأمر لعد القيم المتشابهه.

مثال ٢٧: يمثل الشكل(٣٢-٣) حضور الطلبة في فصل صيفي، فإذا كان الرقم واحد يمثل حضور الطالب وصفر يمثل غيابه، لحساب عدد حضور كل طالب طوال الفصل نقوم بما يلي:

١) بعد النقر على Count من قائمة Transformation تظهر لك نافذة Count Occurrences of values كما في الشكل (٣٣-٣)، نقوم بإدخال اسم متغير جديد يمثل عدد مرات الحضور Presence في مربع Target Variable.

٢) بواسطة سهم الإضافة أنقل القيم lecture1 إلى lecture8 إلى مربع Numeric Variables.

٣) إضغط على زر Define Values فتظهر لك نافذة Count Values within cases المبينة بالشكل(٣٤-٣).

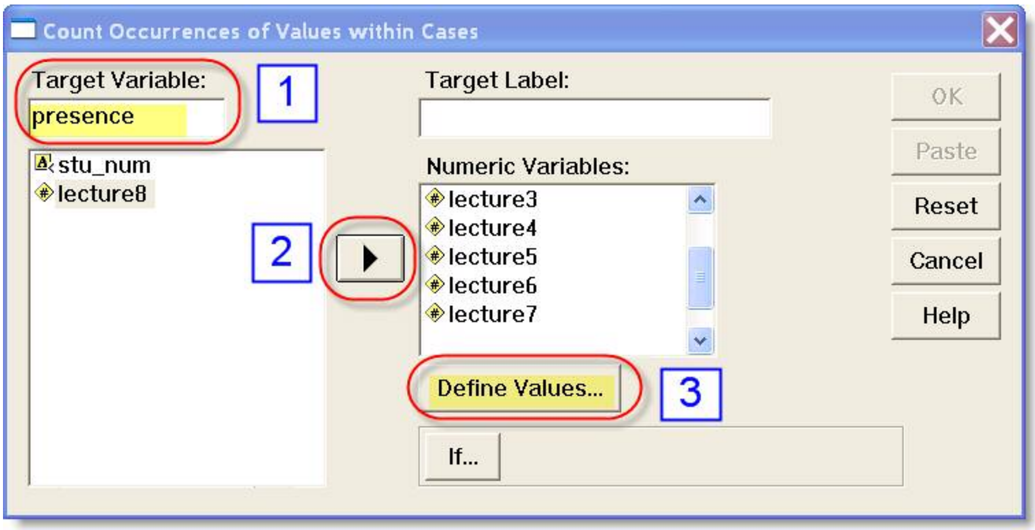
٤) أدخل القيمة التي على أساسها سيتم العد في مربع Value، وفي هذا المثال تساوي واحد.

	stu_num	lecture1	lecture2	lecture3	lecture4	lecture5	lecture6	lecture7	lecture8
1	Omar	1	0	1	0	1	1	0	0
2	Nasser	1	1	0	1	0	1	1	1
3	Ahmed	1	1	1	1	1	1	1	1
4	Raed	1	0	0	0	1	0	1	1
5	Zaid	1	0	1	0	1	1	0	1
6	Amjed	1	1	1	0	1	1	1	0

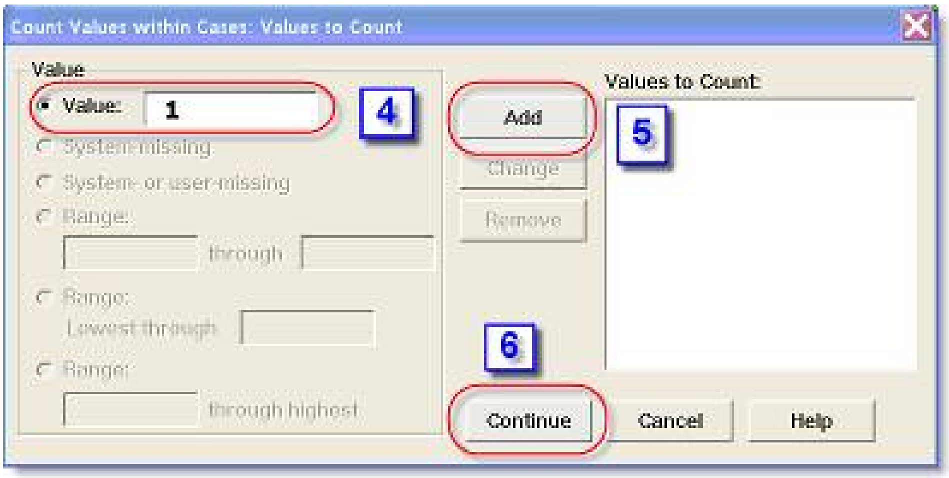
الشكل(٣٢-٣)

٥) أنقر على زر Add.

٦) أنقر على زر Continue ثم Ok فيظهر عدد مرات حضور كل طالب في العمود الأخير كما في الشكل (٣-٣٥).



الشكل(٣-٣٣) نافذة Count Occurrences of values

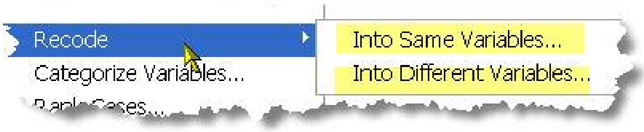


الشكل(٣-٣٤) نافذة Count Values within cases

	stu_num	lecture1	lecture2	lecture3	lecture4	lecture5	lecture6	lecture7	lecture8	presence
1	Omar	1	0	1	0	1	1	0	0	4
2	Nasser	1	1	0	1	0	1	1	1	6
3	Ahmed	1	1	1	1	1	1	1	1	8
4	Raed	1	0	0	0	1	0	1	1	4
5	Zaid	1	0	1	0	1	1	0	1	5
6	Amjed	1	1	1	0	1	1	1	0	6

الشكل(٣-٣٥) نتيجة تطبيق الأمر Count

➢ ٣,١١,٣ الأمر إعادة الترميز **Recode**.



في بعض المسائل قد نقسم العينة إلى مجموعــات ، كتقســيم العلامات إلى فئات و من طرق الترميز :

✓ إعادة الترميز باستخدام متغير جديد Recode into Different Variables.

ننشئ هنا متغير جديد كما في المرات السابقة يحتوي هذا المتغير على رمز المجموعة.

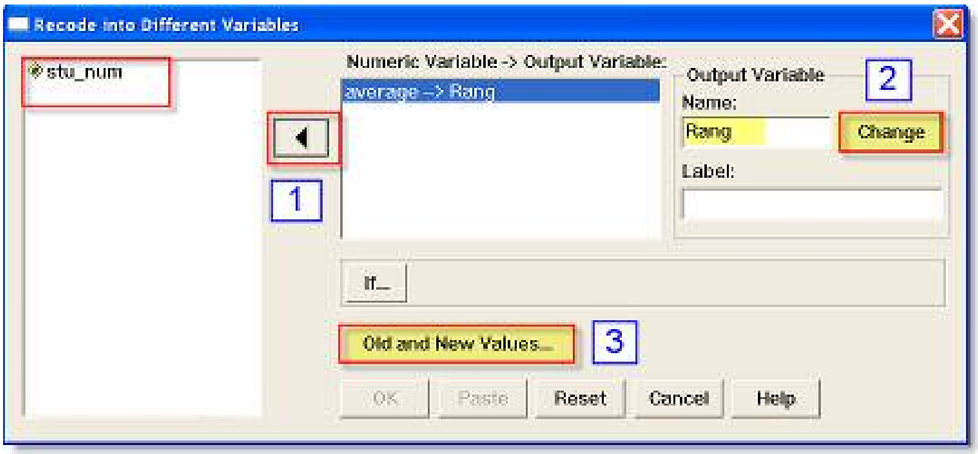
مثال ٢٨: إذا أردت أن نقسم المعدل التراكمي إلى فئات حسب التقسيم المبين أدناه نتبع ما يلي::

المجموعة	المعدل
: A	90 -100
: B+	85 – 89
: B	80- 84
: B-	75 – 79
: C	70 – 74
: D	60 – 69

١. من قائمة الفرعية Recode أنقر على into Different Variables ، عندها تظهر لك نافذة Recode into Different Variables كما في الشكل (٣٦-٣)، إنتق Average من قائمة المتغيرات ثم إضغط على السهم المجاور.

٢. أكتب اسم المتغير الجديد Rang في مربع Name ثم أنقر على Change.

٣. أنقر الآن على زر Old and New Value، سيتم فتح نافذة جديدة كما في الشكل (٣٨-٣).



الشكل (٣٦-٣)

٤. أدخل في الجزء old value الفترة الأولى من ٩٠ إلى ١٠٠ في المربع Range.

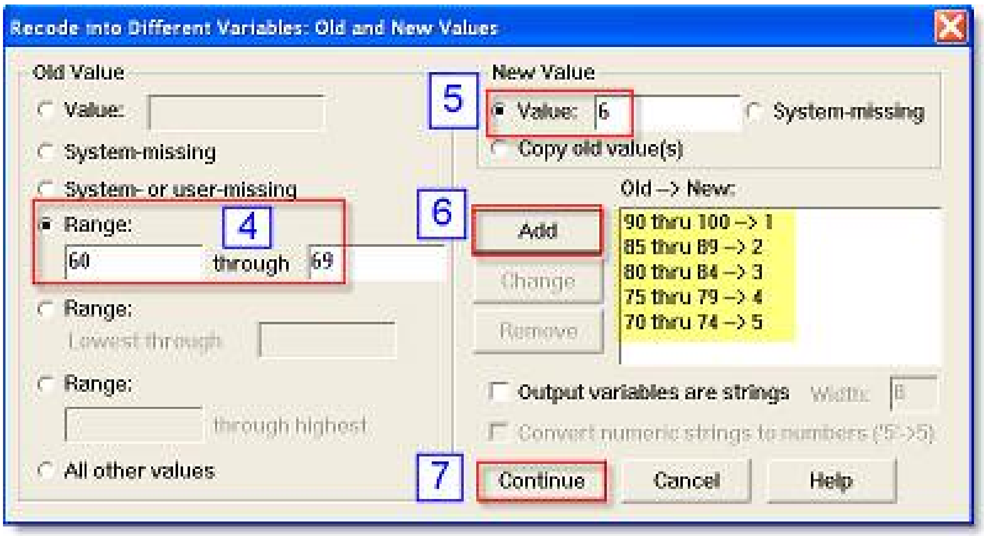
٥. في الجزء new value أدخل في مربع Value القيمة الممثلة لهذه الفترة ولتكن واحد.

٦. ثم إضغط Add لإضافة هذه الفترة إلى القائمة.

٧. كرّر هذه الخطوات لبقية الفترات ثم إضغط Continue ثم من النافذة السابقة إضغط Ok. ستظهر النتيجة كما في الشكل (٣٧-٣).

	stu_num	average	rang
1	200512	88.00	2
2	200515	82.40	3
3	200488	78.20	4
4	200189	72.00	5
5	200290	75.00	4
6	200452	72.50	5
7	200419	88.60	2
8	200333	90.80	1
9	200261	82.00	3
10	200125	73.00	5
11	200122	77.00	4

الشكل(٣٧-٣)



الشكل(٣٨-٣)

☑ ٣,١٢ قائمة Analysis:

هناك عدة أنواع من المتغيرات يتم التعامل معها في SPSS هي:

المتغيرات الاسمية Nominal Variables: وهي متغيرات لا يتم إجراء العمليات الحسابية عليها لعدم وجود قيم لها وتتكون من عدد محدد من الفئات، ولا يوجد اختلاف بين هذه الفئات ، مثل متغير "الجنس" الذي يتكون من فئتين : ذكور وإناث.

المتغيرات الترتيبية Ordinal Variables: يوجد قيمة ترتيبية لهذه المتغيرات التي تتكون من عدد محدد من الفئات، هنا يمكن التميز بين هذه الفئات وترتيبها، ولكن لا يمكن تحديد الفرق بينها، ومتغير العمر مثال على هذا النوع فنقول الطفل الأول اصغر من الثاني ولكن لا تستطيع تحديد هذا الفرق.

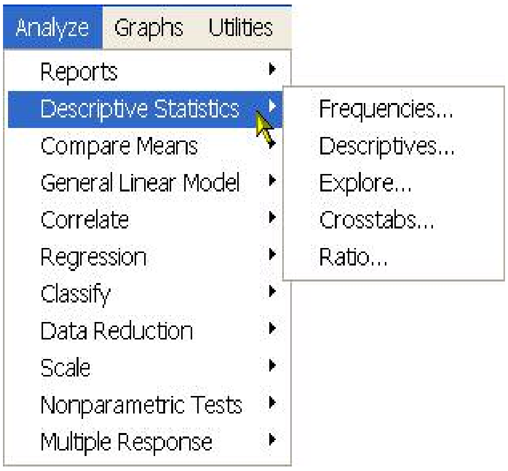
المتغيرات الفئوية Interval Variables: هي المتغيرات التي تستطيع إجراء العمليات الحسابية عليها لأنها تملك قيمة محددة و تستطيع أن تحدد الفرق بين أي متغيرين من نفس الفئة بدقة.

تحتوي قائمة التحليل Analysis على عدة أوامر منها:

➢ **3.12.1 Descriptive Statistics:** تحتوي هذه القائمة الفرعية على عدة أوامر كما في الشكل (٣٩-٣) أولها :

▪ 3.12.1.1 Frequencies: يستخدم هذا الأمر لوصف أفراد العينة وكذلك يستخدم

- لعمل الرسوم البيانية مثل Bar Chart و Pie Chart والتي تستخدم في تمثيل التكرارات و النسب المئوية في حالة المتغيرات النوعية أو الترتيبية وستخدم أيضا لعمل Histogram من اجل المتغيرات الكمية،
- كما يمكن من خلاله إيجاد مقاييس النزعة المركزية Central Tendency من وسط حسابي Mean، وسيط Median، منوال Mode ومجموع Sum

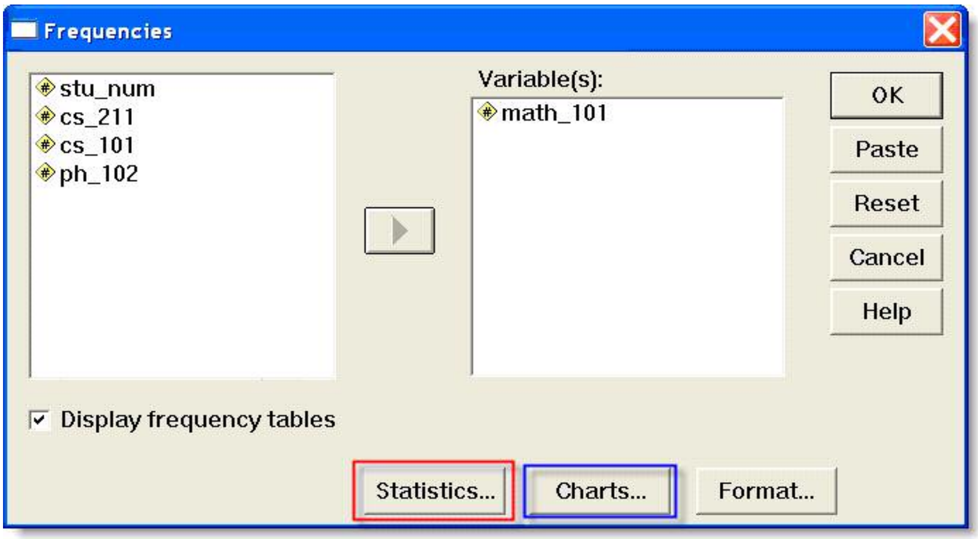


الشكل(٣٩-٣)

- وكذلك يمكن إيجاد مقاييس التشتت كالانحراف المعياري Std Deviation ، التباين Variance، المدى Rang والخطاء المعياري S.E mean والقيمة العظمى Max والصغرى Min.
- وكذلك لإيجاد المئينات Percentiles والرُبّيعات Quartiles
- والإحصاءات الخاصة بشكل التوزيع من التواء Skewness و تفلطح Kurtosis

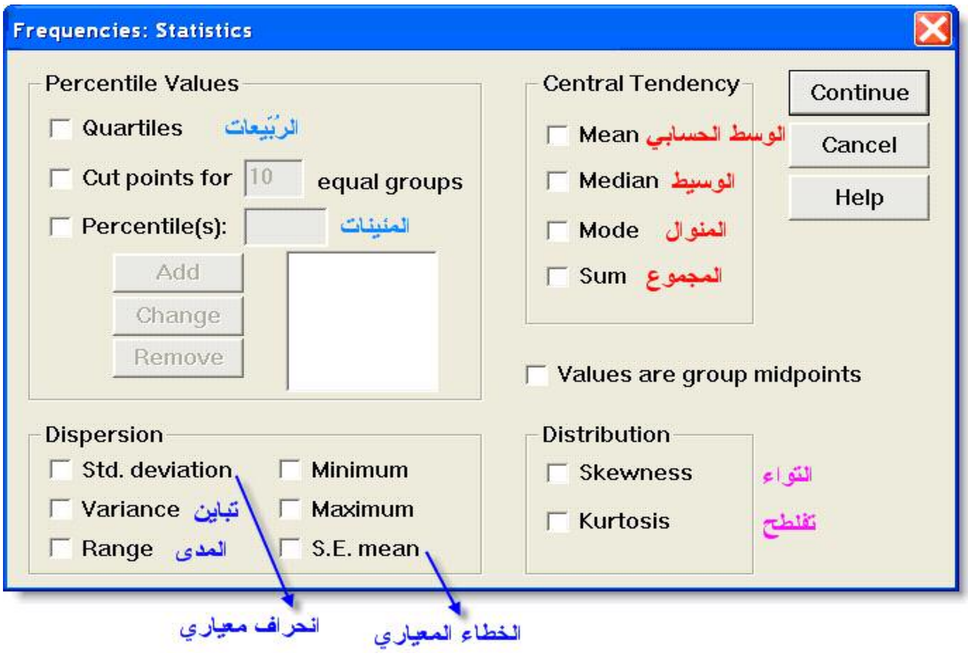
ولحساب هذه الإحصاءات لمتغير نقوم بما يلي:

١. من Analysis أنقر على Descriptive Statistics ثم إنتق منها Frequencies فتظهر لك نافذة كما في الشكل(٤٠-٣) تحدد منها أسماء المتغيرات المطلوبة وذلك بإضافتها إلى مربع Varaiable(s).



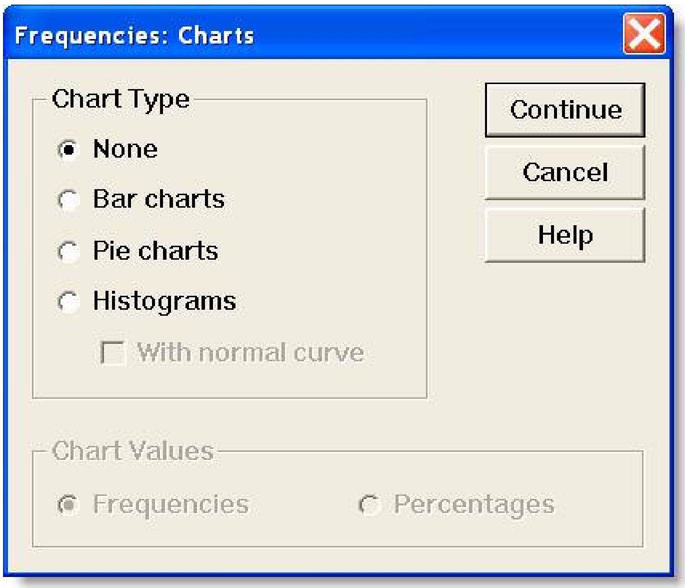
الشكل(٤٠-٣)

٢. إضغط على statistics فتظهر نافذة جديدة كما في الشكل (٤١-٣) إنتق منها ما نريد إيجاده من إحصاءات لهذه المتغيرات ثم إضغط Continue.



الشكل(٤١-٣)

٣. ومن اجل الرسوم البيانية إضغط من النافذة الأولى(٤٠-٣) على Charts فتظهر النافذة التالية (٤٢-٣) تحدد من هذه النافذة الرسم المطلوب ثم إضغط Continue.



الشكل(٤٢-٣)

٤. إضغط على Ok فتظهر لك النتيجة كما في الشكل(٤٣-٣).

Frequencies

Statistics

MATH_101

N	Valid	6
	Missing	0
Mean		75.50
Std. Error of Mean		3.658
Median		77.50
Mode		60[a]
Std. Deviation		8.961
Variance		80.300
Skewness		-1.111
Std. Error of Skewness		.845
Kurtosis		1.143
Std. Error of Kurtosis		1.741
Range		25
Minimum		60
Maximum		85
Sum		453
Percentiles	25	68.25
	50	77.50
	75	82.75

a. Multiple modes exist. The smallest value is shown

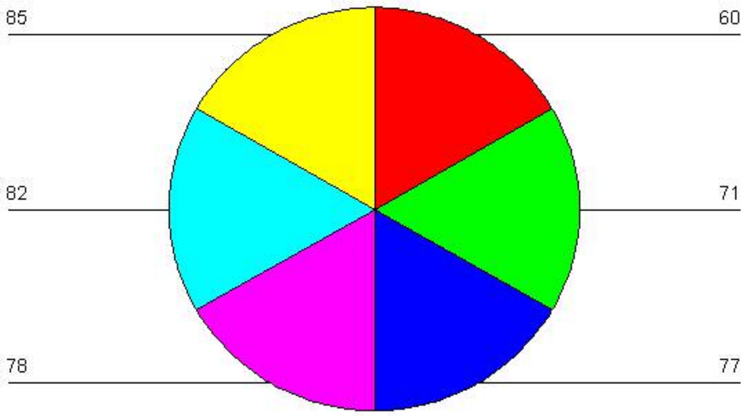
MATH_101

		التكرار Frequency	النسبة المئوية Percent	Valid Percent	Cumulative Percent
Valid	60	1	16.7	16.7	16.7
	71	1	16.7	16.7	33.3
	77	1	16.7	16.7	50.0
	78	1	16.7	16.7	66.7
	82	1	16.7	16.7	83.3
	85	1	16.7	16.7	100.0
	Total	6	100.0	100.0	

النسبة التراكمية في المتغيرات الغير نوعية

النسبة المئوية بعد استبعاد القيم المفقودة

Math_101 للمتغير Pie Chart



الشكل(٣-٤٣) نتيجة استخدام الأمر Frequencies

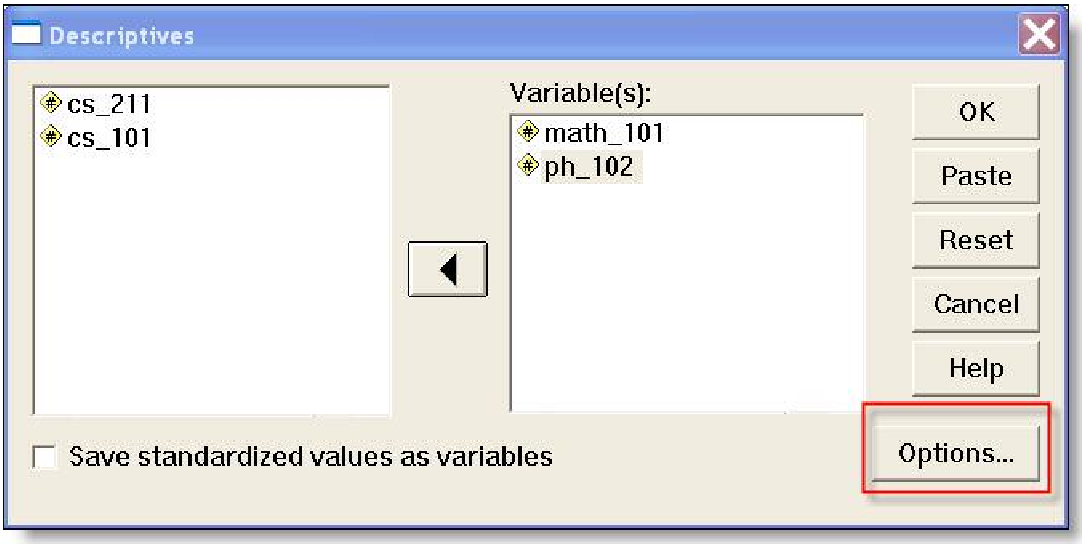
من خلال النتيجة نرى أن العينة تتكون من ٦ علامات أعلاها ٨٥ واقلها ٦٠ بوسط حسابي قدره ٧٥,٥ وانحراف معياري ٨٠,٣ ومن المئينات في القسم السفلي نجد أن ٢٥% من العلامات اقل من ٦٨,٢٥ وان ٥٠% من العلامات اقل من ٧٧,٥ و ٧٥% من العلامات اقل من ٨٢,٧٥.

- 3.12.1.2 Descriptive: يستخدم هذا الأمر مع المتغيرات الكمية فقط.

١. من Analysis أنقر على Statistics Descriptive ثم إنتق منها Descriptive فتظهر لك نافذة كما في الشكل(٣-٤٤) تحدد منها أسماء المتغيرات المطلوبة وذلك بإضافتها إلى مربع Varaiable(s).

٢. إضغط على Options فتظهر نافذة جديدة كما في الشكل (٣-٤٥) إنتق منها ما نريد إيجاده من إحصاءات لهذه المتغيرات ثم إضغط Continue.

٣. إضغط على Ok فتظهر لك النتيجة كما في الشكل(٣-٤٦).



الشكل(٤٤-٣)

الشكل(٤٥-٣)

Descriptives

Descriptive Statistics

عدد المفردات

اسماء المتغيرات	N	Range	الصغرى Minimum	العظمى Maximum	Sum	الوسط Mean	Std. Deviation	Variance
MATH_101	6	25	60	85	453	75.50	8.961	80.300
PH_102	6	18.00	70.00	88.00	464.00	77.3333	5.98888	35.867
Valid N (listwise)	6							

الشكل(٤٦-٣) نتيجة استخدام الأمر Descriptive

النتيجة : جمعت علامات مادة الرياضيات والفيزياء التراكمية لستة طلاب فتراوحت علامات الطلاب في مادة الرياضيات بين ٦٠ و ٨٥ بوسط حسابي قدره ٧٥,٥ وانحراف معياري ٨٠,٣ كما تراوحت العلامات لمادة الفيزياء بين ٧٠ إلى ٨٨ بوسط حسابي قدره ٧٧,٣ وانحراف معياري ٣٥,٨٦.

➢ 3.12.2 Explore:

- يستخدم هذا الأمر لفحص البيانات وتصحيح الأخطاء كوجود قيم شاذة، فترات انقطاع أو وجود قيم غير منطقية.
- يستخدم لتأكد من تحقق بعض الشروط قبل إجراء بعض الاختبارات الإحصائية كتحليل الانحدار والتباين الذي يتطلب أن يكون التوزيع طبيعي.
- و يستخدم للمقارنة توزيع متغير لمجموعتين من الأفراد.
- حساب مقاييس النزعة المركزية ومقاييس التشتت مثل M-Estimators.

عمل بعض الرسومات التي توضح شكل التوزيع مثل Box Plot, Stem and Leaf Plot, Histograms.

☑ ٣,١٣ اختبار الفرضيات

➢ ٣,١٣,١ اختبار T .

☑ **اختبار T** : يستخدم هذا الاختبار لفحص الوسط الحسابي لفرضية ما. ويستخدم هذا الاختبار في حالة

- عدم اعتماد البيانات على بعضها.
- وان يكون توزيع العينة طبيعي إذا كان حجمها صغير وإذا كان حجمها كبير فلا يهم شكل التوزيع.

ومن أنواع اختبارات T :

✓ اختبار T للعينة الواحدة One Sample T-Test.

- يستخدم لفحص وسط عينة والتي تساوي قيمة ثابتة بناءا على الفرضية التالية $H_o : \mu = m$ حيث m قيمة ثابتة تحدد من خلال الإجابات التي جمعت من الاستبيان.
- ليكن x هو المتغير المراد اختبار وسطه يساوي قيمة ثابتة أم لا وذلك من خلال المعادلة التالية :

$t = \dfrac{\bar{X} - m}{\sigma/\sqrt{N}}$ حيث قيمة t تعني عدد الانحرافات المعيارية $\sigma/\sqrt{N}$ الموجودة في الفرق بين الوسط الحسابي والقيمة $\bar{X} - m$ فإذا كانت قيمة t تساوي صفر فان قيمة الوسط الحسابي تساوي m ويزداد الفرق بين الوسط و m كلما ابتعدت قيمة t عن الصفر بالاتجاه الموجب أو السالب.

- لإجراء الاختبار على عينة حسب الفرضية العدمية : لا توجد علاقة سلبية بين افتقار المدارس لوسائل التكنولوجيا الحديثة وبين انخفاض قدرة الطلبة على التحصيل والإبداع. إنتق $\mu = 2$ على فرض أن الإجابة على الاستبيان ستكون الخيار الثاني (الإجابة موافق).

$$H_o : \mu = 2$$

نختبر عند مستوى الثقة confidence level = 95%

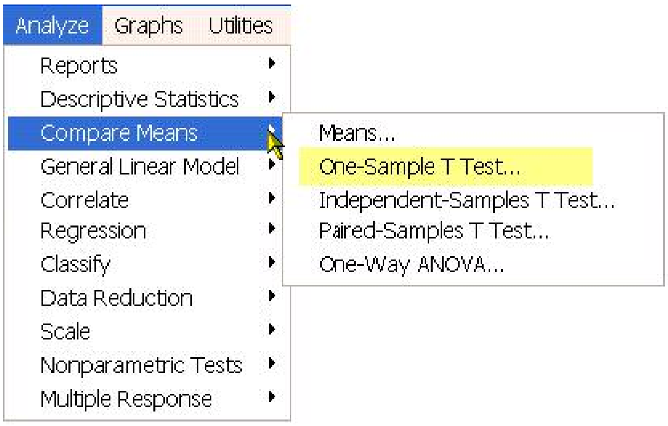
و مستوى معنوى significance level = 5% (احتمال وجود خطأ $\alpha = 0.05$)

مقياس لكر :

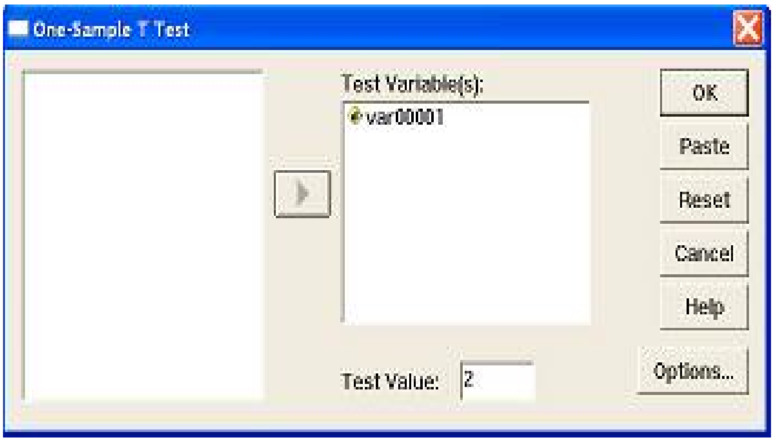
موافق بشدة(١)	موافق(٢)	محايد(٣)	ارفض(٤)	ارفض بشدة(٥)

لحساب نتيجة الاختبار نقوم بما يلي:

١. أنقر على One Sample T Test من قائمة Compare Means الفرعية كما في الشكل(٤٧-٣).
٢. تظهر لك النافذة المبينة بالشكل (٤٨-٣)، أضف المتغير المراد اختباره إلى قائمة Variables .
٣. إطبع قيمة الثابت في Test Value وفي هذا المثال تساوي ٢.
٤. إضغط Ok.



الشكل(٤٧-٣)



الشكل(٤٨-٣)

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
VAR00001	18	2.1872	.06238	.01470

One-Sample Test

	Test Value = 2					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
VAR00001	12.733	17	.000	.1872	.1562	.2182

الشكل(٤٩-٣) نتائج اختبار T للعينة الواحدة

النتيجة: يبين الشكل (٤٩-٣) قيمة الوسط الحسابي Mean والانحراف المعياري والخطأ المعياري للمتغير الذي اخترناه سابقا لفحص وسطه الحسابي، كما يبين أيضا متوسط الفرق بين المتغير والقيمة المفترضة Mean Difference وكانت تساوي في هذا المثال ٠,١٨٧، تستطيع الإجابة على الفرضية السابقة " لا توجد علاقة سلبية بين افتقار المدارس لوسائل التكنولوجيا الحديثة وبين انخفاض قدرة الطلبة على التحصيل والإبداع" من خلال مقارنة قيمة sig.(2-tailed) ويطلق عليها أيضا P-value مع قيمة مستوى المعنوية significance level وهو غالبا = 5% $(\alpha = 0.05)$ فإذا كانت $(P\text{-value} \rangle \alpha)$ فإننا نقبل الفرضية والاّ نقوم برفضها وفي هذا المثال قيمة P-value هي صفر و α هي ٠,٠٥ أي أن P اصغر وبالتالي نرفض الفرضية السابقة وبالتالي تكون هناك علاقة بين افتقار المدارس إلى التكنولوجيا وانخفاض التحصيل.

سوف نناقش العلاقة الخطية بين المتغير الكمي التابع والمستقل من خلال الارتباط الخطي Linear Correlation و تحليل الانحدار الخطي Linear Regression .

☑ ٣,١٤ الارتبـــاط Correlation

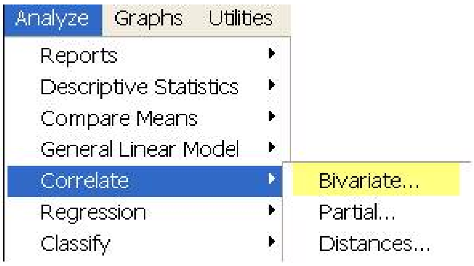
يستخدم لتحديد قوة واتجاه العلاقة بين متغيرين كميين، وقد تكون النتيجة غير دقيقة لتأثر المتغير التابع بعوامل أخـــرى غير العامل المستقل لهذا يجب عزل المتغير التابع عن باقي العوامل إذا أمكن ونستخدم هنا الارتباط الخطـــي الجزئـــي Partial Linear Correlation بعد استبعاد متغير أو اكثر. كما نستخدم ارتبـــاط بيرســـون Pearson Correlation Coefficient مع المتغيرات الترتيبية Ordinal ومن خلال الاختبار الإحصائي تحدد وجود أو عدم وجود ارتباط بـــين المتغيرين وحجمه، حتى تستطيع أن نستخدم النوع الأخير الارتباط يجب أن يكون توزيع كل متغير يدخل الاختبار توزيع طبيعي وقيم المتغيرين لشخص ما مستقلة وفي حالة عدم تحقق الشرطين نغير نوع الارتباط إلى سبيرمان أو كاندال تـــاو ب Kendall Tau-b.وتقع قيمة الارتباط بين ١− و ١+.

مثال ٢٩: حدد نوع العلاقة وقوتها بين المتغيرين رأس المال و الربح من ملف data.

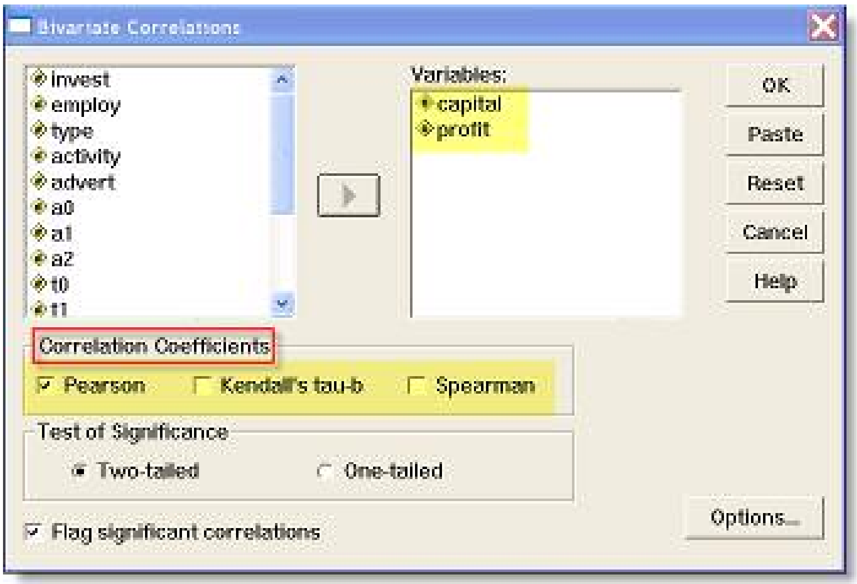
الحل:

١. بعد فتح ملف البيانات أنقر على analyze ثـــم أنقـــر علـــى Correlation ونختار منه Bivariate كما في الشكل(٥٠-٣)، ستظهر لك النافذة المبينة بالشكل(٥١-٣).

٢. إنتق اثنين أو اكثر من المتغيرات المراد إيجاد ارتباطها ونضيفها إلى مربع variables.



الشكل(٥٠-٣)



الشكل(٥١-٣)

٣. إنتق نوع الارتباط المطلوب وهو هنا معامل ارتباط بيرسون .

٤. إنتق الآن Ok فتظهر النتيجة كما في الشكل(٥٢-٣).

Correlations

Correlations

		CAPITAL	PROFIT
CAPITAL	Pearson Correlation	1	-.029
	Sig. (2-tailed)	.	.774
	N	100	100
PROFIT	Pearson Correlation	-.029	1
	Sig. (2-tailed)	.774	.
	N	100	100

الشكل(٣-٥٢) النتيجة

النتيجة: تبين النتيجة حجم العينة N وعددها ١٠٠ عنصر، و الارتباط الثنائي بين كل زوج من المتغيرات. وتم استخراج ارتباط بيرسون لفحص وجود علاقة بين حجم رأس المال و الربح، وقد وجد من خلال النتائج في الشكل(٣-٥٢) أن هناك علاقــة ذات دلالة إحصائية بين capital و profit وكانت هذه العلاقة موجبة طرديه حيث بلغ معامل الارتباط ٠,٧٧٤ وتعتبر علاقة قوية.

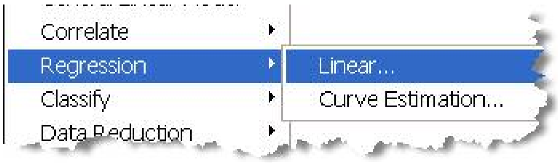
☑ ٣,١٥ الانحدار البسيط Linear Regression:

يستخدم تحليل الانحدار الخطي Regression Linear لتمثل العلاقة بين متغيرين على شكل معادلة خطية بهدف التنبؤ بقيمة المتغير التابع إذا عرفت قيم المتغير المستقل وتكتب المعادلة على الشكل $Y = \alpha + \beta X + e$ حيث Y المتغير التابع الذي نريد التنبؤ بقيمته و X المتغير المستقل و e الخطأ في التقدير.

مثال ٣٠: أوجد معادلة الانحدار الخطي بين الطول والوزن لمجموعة من الطلاب.

الحل:

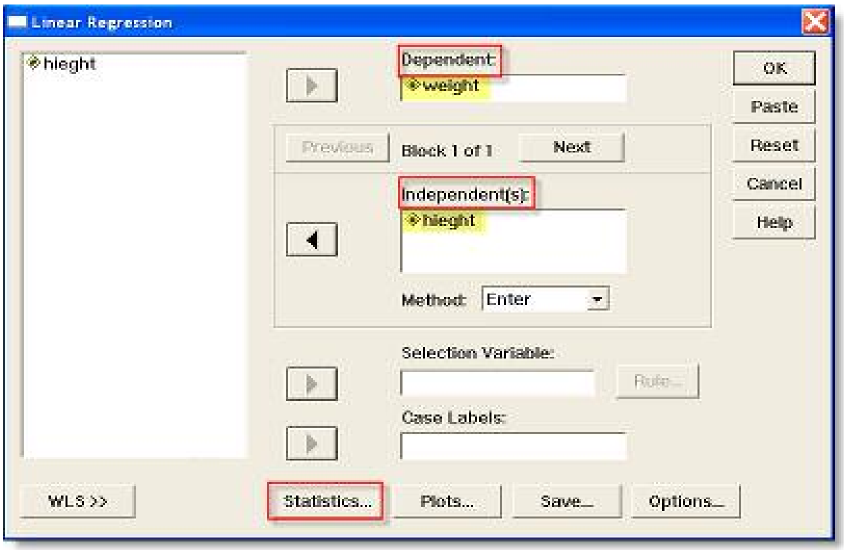
١. بعد فتح ملف البيانات أنقر على analyze ثم أنقر على Regression ونختار منه Linear ستظهر لك النافذة المبينة بالشكل(٣-٥٤).



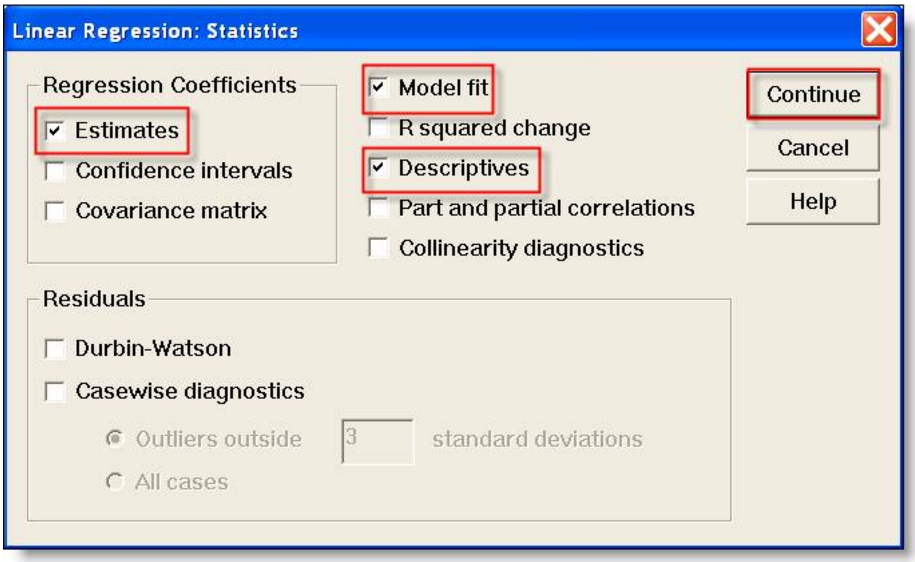
الشكل(٣-٥٣)

٢. إنتق المتغير المستقل height ونضيفه إلى مربع independent و أضف المتغير التابع weight إلى مربع dependent.

٣. أنقر على statistical فتظهر لك نافذة Linear Regression: statistical كما في الشكل(٣-٥٥)



الشكل(٣-٥٤)



الشكل(٤٥-٣)

٤. إنتق Descriptive ونتأكد من اختيار Estimate و Model Fit .

٥. إضغط Continue ثم Ok، فتظهر النتائج كما في الشكل(٥٦-٣).

Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	HEIGHT[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: WEIGHT

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.895[a]	.801	.768	2.4424

a. Predictors: (Constant), HEIGHT

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1 -37.746	20.407		-1.850	.114
	HEIGHT	2 .608	.124	.895	4.917	.003

a. Dependent Variable: WEIGHT

الشكل(٥٦-٣) نتيجة تحليل الانحدار

النتيجة: كما يظهر في الشكل(٥٦-٣) تم حساب قيمة R^2 البالغة ٠,٨٠١ التي تدل على قدرة متغير الطول في التنبؤ في الوزن.وهي مربع معامل الارتباط وهو ارتباط طردي موجب قوي. ثم حسبت قيمة α و β في جدول Coefficients وهي تدل على شكل معادلة الانحدار (التنبؤ) على الشكل $weight = -37.746 + 0.608 \times height$ من الشكل رقم واحد يمثل α و رقم اثنان يمثل β.